2021-06

PRIVACY RISK ASSESSMENT OF A DON DIGITAL CONTACT TRACING SYSTEM USING THE NIST PRIVACY FRAMEWORK

Carter, Thomas E.

Monterey, CA; Naval Postgraduate School

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

PRIVACY RISK ASSESSMENT OF A DON DIGITAL CONTACT TRACING SYSTEM USING THE NIST PRIVACY FRAMEWORK

by

Thomas E. Carter

June 2021

Thesis Advisor: James B. Michael
Co-Advisor: Joshua A. Kroll

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** June 2021	**3. REPORT TYPE AND DATES COVERED** Master's thesis

4. TITLE AND SUBTITLE
PRIVACY RISK ASSESSMENT OF A DON DIGITAL CONTACT TRACING SYSTEM USING THE NIST PRIVACY FRAMEWORK

5. FUNDING NUMBERS

6. AUTHOR(S) Thomas E. Carter

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
N/A

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release. Distribution is unlimited.

12b. DISTRIBUTION CODE
A

13. ABSTRACT (maximum 200 words)

COVID-19 has impacted the DON's readiness and ability to operate effectively, and this presents a potential security risk to the U.S. population. In order to allow employees to return to their normal working routines and prevent the spread of COVID-19, the DON began to procure and test a Bluetooth-based contact tracing system in 2020. This research explores the privacy considerations of a digital contact tracing system that was being procured by the DON, and it does so by applying the National Institute of Standards and Technology (NIST) Privacy Framework which was released in January 2020. We are not only able to provide recommendations about the privacy of the contact tracing system, but we are also able to assess the privacy framework as a privacy risk management tool.

We provide a privacy threat model of the system by analyzing the data path of the contact tracing system. We also apply the NIST Privacy Framework to our model of the system, and we determine that the framework is useful for risk identification but does very little to contribute to assessing the impact or likelihood of privacy risks. The threat modeling also reveals that the DON needs to focus more on disassociability of data sets when considering the privacy risks of the contact tracing system, and we recommend that the DON begin conducting privacy testing on systems that collect PII. Finally, we recommend combining the NIST Cybersecurity and Privacy Frameworks in order to streamline the assessment process.

14. SUBJECT TERMS privacy, contact tracing, risk management framework, centralized, decentralized, National Institute of Standards and Technology, NIST, NIST Privacy Framework, threat modeling, system engineering, deidentification, disassociability	**15. NUMBER OF PAGES** 149
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

PRIVACY RISK ASSESSMENT OF A DON DIGITAL CONTACT TRACING SYSTEM USING THE NIST PRIVACY FRAMEWORK

Thomas E. Carter
Lieutenant, United States Navy
BS, University of Maryland College Park, 2014

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN COMPUTER SCIENCE

from the

NAVAL POSTGRADUATE SCHOOL
June 2021

Approved by: James B. Michael
Advisor

Joshua A. Kroll
Co-Advisor

Gurminder Singh
Chair, Department of Computer Science

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

COVID-19 has impacted the DON's readiness and ability to operate effectively, and this presents a potential security risk to the U.S. population. In order to allow employees to return to their normal working routines and prevent the spread of COVID-19, the DON began to procure and test a Bluetooth-based contact tracing system in 2020. This research explores the privacy considerations of a digital contact tracing system that was being procured by the DON, and it does so by applying the National Institute of Standards and Technology (NIST) Privacy Framework which was released in January 2020. We are not only able to provide recommendations about the privacy of the contact tracing system, but we are also able to assess the privacy framework as a privacy risk management tool.

We provide a privacy threat model of the system by analyzing the data path of the contact tracing system. We also apply the NIST Privacy Framework to our model of the system, and we determine that the framework is useful for risk identification but does very little to contribute to assessing the impact or likelihood of privacy risks. The threat modeling also reveals that the DON needs to focus more on disassociability of data sets when considering the privacy risks of the contact tracing system, and we recommend that the DON begin conducting privacy testing on systems that collect PII. Finally, we recommend combining the NIST Cybersecurity and Privacy Frameworks in order to streamline the assessment process.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AI	artificial intelligence
ATO	authority to operate
BLE	Bluetooth low energy
BUMED	Bureau of Medicine and Surgery
CIA	confidentiality, integrity, availability
CIO	Chief Information Officer
COTS	commercial off-the-shelf
CPTM	Cloud Privacy Threat Modeling
CSAIL	Computer Science and Artificial Intelligence Laboratory
CSG	Carrier Strike Group
DFD	data flow diagram
DOD	Department of Defense
DON	Department of the Navy
DP-3T	Decentralized Privacy-Preserving Proximity Tracing
DPD	Data Protection Direction
GDPR	General Data Protection Regulation
GPS	Global Positioning System
HIPAA	Health Insurance Portability and Accountability Act
IOT	Internet of Things
LINDDUN	linkability, identifiability, non-repudiation, detectability, disclosure of information, unawareness, non-compliance
MHS	Military Health System
MIT	Massachusetts Institute of Technology
ML	machine learning
NIST	National Institute of Standards and Technology
NISTIR	NIST Internal Report
NSWC	Naval Surface Warfare Center
OPNAV	Office of the Chief of Naval Operations
OPSEC	Operational Security
PACT	Private Automated Contact Tracing
PEPP-PT	Pan-European Privacy Preserving Proximity Tracing

PHI	protected health information
PIA	Privacy Impact Assessment
PII	personally identifiable information
RFI	request for information
RMF	Risk Management Framework
SDLC	System Development Life Cycle
SP	Special Publication
STRIDE	spoofing, tampering, repudiation, information disclosure, denial of service, escalation of privilege
USNA	United States Naval Academy

ACKNOWLEDGMENTS

I would like to acknowledge the work being done on a Department of the Navy contact tracing system by the engineers at Naval Surface Warfare Center Crane Division. With their help, I was able to glean valuable details about the procurement and assessment of commercial-off-the-shelf (COTS) contact tracing systems and some of the impediments to implementation in the fleet.

I would also like to acknowledge the work being done by OPNAV N09D (Digital Transformation Office) toward integrating COVID-19 data into the Jupiter system. This work provided me valuable insight when modeling the backend data flow of a potential DON contact tracing system.

Finally, I want to thank my thesis advisors, Dr. Bret Michael and Dr. Joshua Kroll. In addition to our invaluable discussions and their critical feedback, their greatest contribution to the research was how they fundamentally shaped my approach to solving complex issues such as the one addressed in this thesis. They taught me how to think about problems the way a computer scientist would and how to tackle real-world problems that span multiple domains of expertise.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

Privacy surrounding digital data and use of technologies has been a topic of concern at the federal level in the United States for many decades, as evidenced by watershed events such as the enactment of the Privacy Act of 1974. As artificial intelligence (AI), machine learning (ML), and big data analytics continue to advance, digital privacy and privacy-preserving technologies have become ever more prominent in local, national, and international discourse. The federal government's current work in the digital privacy space centers around the National Institute of Standards and Technology (NIST) Privacy Framework, which was released in January 2020 [1]. NIST created the framework to help organizations identify privacy-related risk. Such risk arises during the development and sustainment of applications and systems. The hope is that applying the framework will foster a dialog about privacy, privacy-related risks, and the management of those risks at all levels of development and sustainment.

For the purpose of this work, we are using the NIST definition of privacy as articulated in the NIST Privacy Framework. NIST describes privacy as a "concept that helps to safeguard important values such as human autonomy and dignity," but the framework also acknowledges the shifting nature of privacy and how it adds to the difficulty in communicating privacy risks [1]. In this thesis, we use that broad concept of privacy to extend the definition to include safeguarding important values of organizations as well. This is important since the DON is concerned with the organizational as well as individual privacy of its employees. An in depth discussion about the definition of privacy is outside of the scope of this thesis, but for a more detailed analysis of debate around the definition of privacy, see Mulligan et al. [2].

This research presents an application of the NIST Privacy Framework to a conceptual model of a Bluetooth-based proximity tracing system inspired by infectious diseases such as COVID-19. Our model is designed to capture key characteristics of a system the Department of the Navy (DON) plans to acquire. We understand that the concept of "contact tracing" is not always synonymous with proximity tracing since contact tracing usually involves the work of other entities (i.e., medical professionals) in addition

to any proximity tracing data and devices that are used. For the rest of this thesis, we refer to the DON system as a contact tracing system since this is the accepted language in scientific literature for this type of system, and our design of the system also includes more components than just the proximity tracing devices.

The digital contact tracing system described in this thesis collects contact tracing data which then flows to a back-end system called Jupiter (which is the current DON name for its newly created enterprise data management system). Within Jupiter, the DON will be able to perform analysis on the collected contact tracing data while combining it with data from other DON databases. This consolidated analysis will enable the preparation and dissemination of fleet-readiness products for high-level decision makers concerning medical-readiness issues such as those related to COVID-19 cases impacting the DON workforce.

The proposed Navy contact tracing system poses privacy concerns for the DON workforce. Privacy violations can lead to a loss of trust and a loss of security. Privacy is generally considered an essential quality of free and democratic societies, even if the definition of privacy varies among these societies. It is no doubt fundamental to American society. The Department of Defense (DOD) is also concerned with the trustworthiness of systems that process, store, and transmit privacy-related information, both from a technical perspective and how privacy violation could overall affect trust in the DOD and the rest of the federal government.

NIST Internal Report (NISTIR) 8062 defines system trustworthiness as "worthy of being trusted to fulfill whatever critical requirements may be needed for a particular component, subsystem, system, network, application, mission, enterprise, or other entity" while "from a privacy perspective, a trustworthy system is a system that meets specific privacy requirements in addition to meeting other critical requirements" [3]. This means that the concept of privacy needs to be deliberately introduced into information system design. A trustworthy system is currently not assumed to be privacy preserving. Trustworthiness has always been centered on addressing security objectives that do not necessarily encompass privacy.

Loss of trustworthiness of a federal system from a privacy perspective will inevitably bleed over into an individual loss of trust in the DOD enterprise. Just as failing to manage privacy risks has direct adverse consequences on businesses and their employees or customers, an erosion of trust in government institutions will also bring negative effects to employees.

This system may seem like an effort by the federal government to exert more control over its employees, and employees might see this as a means to track their activities. This has civil liberties concerns, especially if the data can easily be compromised. Even worse, it is easy to abuse the data or use it in ways in which it was not meant to be used. For example, what if the government used data from a contact tracing system to determine that a worker was not at his or her duty station during business hours, or what if the data could lead to the revelation that two workers are spending a questionable amount of time in close proximity to each other? This concern could be the result of two different aspects of a proximity tracing system. Either the system could be used to directly track the movements of employees, or private information about employees could be predicted or inferred from aggregated data sets. This is why the privacy controls of this system and the privacy framework by NIST are so important. They enable confidence, and this confidence is part of the foundation for building a capable workforce.

Privacy can also affect physical and operational security. As with all contact tracing systems, individuals can be deanonymized and social graphs can be reconstructed constructed with enough effort from data gathered by the system we consider, and which DON is procuring. It would not be too much of a stretch for a malicious actor to be able to use contact tracing data to be able to reconstruct an organizational chart for a unit, whether that malicious actor is either an authorized user within the system or an outside observer. If this happens, the physical or operational security of the unit is breached. Privacy should be considered an enabler for personnel and unit-level security concerns.

Data assumed to be private could be used by an adversary for general reconnaissance on a military unit. The U.S. military has already seen breaches in private information that can lead to potential physical or operational security threats. In 2018, the military discovered that fitness tracking apps collect a treasure trove of data that is publicly

available, and this data can be used to reveal the locations of military installations overseas. In other instances, the data can be used to identify the movement and identities of other government employees. The U.S. military has confirmed that this discovery has not actually led to a compromise of security, but it is easy to see how this could compromise, at the very least, the physical security of a military unit or operational security of a unit's activities [4].

This fitness-app scenario is just one example of a compromise of publicly available data. There are a growing number of these scenarios, in part driven by the expanding adoption of Internet of Things (IoT) devices. As more data is collected and processed from a myriad of devices (contact tracing systems included), it will only become more apparent that user privacy needs to be taken into account. NIST has already noted public concern over the deployment of smart meters. NIST claims that information collected by smart home devices presents substantial privacy concerns. This includes the possibility of analyzing and tracking individuals or organizations just through their use of smart-energy devices [3].

Privacy and cybersecurity (in the rest of this document, security refers to cybersecurity) are different yet related concepts, and privacy aspects in an information system can and usually do overlap with security concerns. For example, the safe handling of PII is a concern that should be shared by both security engineers and privacy engineers. This concern can arise as the result of unauthorized system behavior (a security concern) or as a byproduct from the authorized processing of PII (a privacy concern) [3]. The bottom line is that privacy and security can affect each other, but these concerns need to be approached differently.

Privacy and security can also be interrelated in the sense that security breaches can lead to privacy compromises, and these privacy compromises can in turn affect national security. The OPM data breaches in 2015 led to a compromise of the private data of millions of Americans [5]. Users of the system assumed the data was secure from unauthorized access. The data turned out to be highly privacy-sensitive in nature. The stolen data could be damaging to individuals (e.g., by exposing them to identity theft), and it could also be used by foreign adversaries to conduct cyber attacks and espionage. This

leads to the question of whether or not this data could have been stored in a way that preserved the privacy of the individuals even in the event of a compromise.

In order to ensure that the DON Contact Tracing System is able to effectively protect the privacy of the members of its workforce, we conducted an assessment of both the proposed contact tracing system and the NIST Privacy Framework. For background information, we have detailed the contact tracing system and the larger privacy concerns surrounding the system in Chapters II and III, respectively. In Chapter IV, we apply the NIST Privacy Framework and model some of the privacy-related risks associated with operation of the system. We start by creating a privacy threat model of the contact tracing system to identify threats specific to the system, and then we apply the NIST Privacy Framework (see Appendix B). Using the NIST Privacy Framework as a guide, we then apply privacy and security controls listed in NIST Special Publication (SP) 800–53 rev5 to the threats that were identified and consider whether they are adequate to controlling privacy risks [6]. We also use the framework to identify new risks or reinforce the severity of known risks. After identifying risks and applying the framework with its controls, we discuss the risk assessment factors of impact and likelihood and how they are necessary to provide a privacy risk assessment to organizations. Then we briefly discuss the extent to which the privacy framework fulfills that role.

In Chapter V, we provide our assessment of the contact tracing system based on the application of the framework. We conclude that management of privacy concerns in this system requires cognizance of the intersection of policy, disassociability, and security when evaluating the degree of privacy afforded by the system. The contact tracing system needs a more robust policy on disassociability since that is a privacy control that is largely disconnected from security. In addition to the privacy policy on disassociability, privacy-related goals of all the stakeholders need to be detailed when designing, testing, or acquiring this contact tracing system. The decision maker and users should also pay more attention to access controls and limit the number of employees that can see re-identified data in the contact tracing data. To assess the system's risk posture with regard to anonymity re-identification, we recommend using a robust, accepted approach such as the HIPAA Expert Determination de-identification methodology. As a further risk control, we recommend that, to the maximum extent possible, only medical professionals are given the

access to identified records that link contact tracing data to names or other personal information as this reinforces the principle of least privilege. Since privacy controls are tailored to specific systems, we also recommend organizations like NSWC Crane implement privacy testing and conduct privacy research in a similar manner to how security testing is done for systems that are procured by the DON. This would streamline the use of the NIST Privacy and Cybersecurity Frameworks in the Risk Management Framework (RMF) process.

Chapter VI provides an assessment of and recommendations for the NIST Privacy Framework. Throughout the thesis we consider the relationship between privacy and security controls and their protentional for contributing to the overall privacy posture of an enterprise. We recommend that NIST consider combining its Privacy and Cybersecurity Frameworks into one framework to enable the streamlining of system assessments. The recommendation assumes that a combined framework will aid users in identifying risks that apply to both privacy and security, but at the same time, aid users in isolating risks that apply just to privacy. We also recommend some constraints and cautions on using the NIST Privacy Framework. It should not be used as a checklist, and its usefulness is highly dependent on the expertise of the organization using it. The organization needs to be able to identify its risk tolerance and also be able determine the likelihood of threats without the framework since the framework does not provide methods or guidance to determine the impact and likelihood of threats. The framework's main utility is in identifying threats and mapping controls to address identified threats. Furthermore, the NIST Privacy Framework is a guide to performing risk assessments. As such, the organization applying the framework needs to be able to go beyond the guidance to carry out the detailed risk assessment of their systems and practices. The thesis concludes with conclusions and recommendations for future research.

II. SYSTEM AND FRAMEWORK DESCRIPTIONS

This chapter provides a description of the high-level architecture of a proposed Bluetooth contact tracing system under consideration for procurement by the Department of the Navy (DON). The system consists of wearable contact tracing devices, storage devices that aggregate data collected about users, and the backend platform that stores and processes the collected data. The backend platform that will receive the data from the storage devices at the various commands is known as Jupiter. Jupiter is the DON's data enterprise management system. Jupiter, via the databases and cloud storage platforms that it hosts, provides for conducting data analytics and, in turn, provides decision makers with situational awareness of fleet readiness. We use this high-level system architecture and the data flow path for contact tracing data as an example application on which to test the NIST Privacy Framework for managing privacy-related risk in the DON.

In addition to providing an overview of the exemplar contact tracing system, this chapter contains a summary of the NIST Privacy Framework. This summary provides necessary background for our assessments of the privacy properties of the example system and the risk management capabilities of the framework itself. The aim of this research is to inform the Navy and other organizations (to include private industries) about the challenges and way ahead for developing, deploying, and sustaining information systems which process, store, and transfer data while effectively managing privacy risks.

A. DEPARTMENT OF THE NAVY CONTACT TRACING SYSTEM

On July 8, 2020, the Department of the Navy issued a Request for Information (RFI) (SAM Notice ID N0002420NR24006) about commercially available proximity tracking systems [7]. This request was released by the Naval Sea Systems Command (NAVSEA) Naval COVID Rapid Response Team (NCR2T) in the wake of the COVID-19 outbreak in early 2020. While the COVID-19 pandemic affected more than just the military, the Navy's need for a technical solution was reinforced by the events of the USS *Theodore Roosevelt* (CVN-71) in which more than 1,200 sailors assigned to the carrier contracted the virus [8]. This demonstrated that the Navy's workforce is vulnerable to outbreaks of communicable diseases such as COVID-19, but more importantly, such

outbreaks can have severe impacts on fleet readiness. Navy leadership concluded that it needs a means for identifying and limiting the spread of COVID-19 outbreak, one of the means being an enterprise-wide contact tracing system.

As for the system itself, the request states that the system components should include "wearable proximity tracking devices" and "storage processing devices" [7]. The users will wear these tracking devices, each of which, identified by a unique identifier, calculates and records the identifiers of and distances from other wearable devices that it comes in proximity to. Distance measurement and identifier exchange is based on Bluetooth signal strength, and the dates/times of the measurements [7].

The system concept is that if User A is in close proximity to User B for an extended period of time, User A's device will collect and store the unique identifier of User B and the time this interaction took place. At periodic intervals, User A's collected unique identifiers and accompanying metadata will be automatically uploaded to a storage processing device (similar to a Wi-Fi router) located on a base or naval vessel. If User B is diagnosed with COVID-19, this diagnosis can be annotated in the system. By scanning the list of every unique identifier that has been in close proximity to User B's unique identifier in the last few weeks, the system can then determine which users to alert, which in this example would be User A. User A would be considered at high risk of having contracted the virus, and User A can choose to get tested and/or quarantine.

The primary benefit of this system is that servicemembers can be identified if they have been in close proximity to other servicemembers that have been positively diagnosed with COVID-19. The secondary benefit, as stated in the request, is that the DON can use this collected data "to determine if social distancing policies put in place by the government employers are effective" at preventing outbreaks of COVID-19 within the workforce [7].

The DON considers the collected proximity data to be important for aggregated use in conducting large-scale analytics. Therefore, an important aspect of contact tracing is how and where the data will be stored after collection, and what policies will be applied to stored contact tracing data. The RFI is does not speak to data storage and administration. When referring to the storage and processing stations, it states that a "Cloud solution is desired that would replicate the functionality of the local, standalone device" and that the

program administrator will "administer any connections to other databases, such as digital medical records" [7]. How the data is collected, stored, processed, and governed within a system has implications for privacy. This is the topic of Chapter III.

After collection, it is reasonable to assume that the proximity data will not remain local to the collecting commands and components; rather, we assume that data will be made available to the wider DON Enterprise Data Management System for processing in order to provide situational awareness to leadership and to facilitate high-level decision-making across the department. The DON likely wants to leverage the power of advanced analytics or AI/ML techniques while also incorporating data from other databases in the DON Enterprise Data Management System in support of this high-level decision-making function.

B. BACKEND DATA FLOW FOR CONTACT TRACING DATA

This section covers the DON's modernization of data management and Jupiter system. This section also describes how the contact tracing system and Jupiter could interoperate.

1. Modernization of DON Data Management

When a proximity contact tracing system collects data from edge devices worn or carried by DON personnel, that data could remain at the local-command level, we assume for the purpose of this research that any proximity data collected by a DON contact tracing system will be made available ultimately to the DON's Enterprise Data Management System, also known as Jupiter, as part of an effort to modernize and optimize the DON's data architecture.

This assumption is based on the posture that the DON has taken around data analytics and data sharing in recent years. There has been a push for modernization and optimization regarding the Navy strategy around data and analytics. In September 2017, the DON chief information officer (CIO) released the "Department of the Navy Strategy for Data and Analytics Optimization" [9]. It describes the vision and goals of the department in regard to data analytics and data management. The strategy centers on the ability to use available department data and analytical assets for the benefit of the DON in

order to "enhance our combat capabilities, increase our operational efficiencies, and improve our ability to make evidence-based decisions quickly" [9].

The DON's position on data management was further codified when it released the "Department of the Navy Information Superiority Vision" in February 2020 [10]. The vision is to "modernize, innovate, and defend" the DON infrastructure; and the document states that the DON currently "lacks a mastery of its Information Environment" [10]. This publication as well as the Navy Strategy publication released in September 2017 emphasize the need to leverage data for decision-making purposes using the most modern means and technology available.

In February 2020, the DON chief data officer provided more specific details about the DON's path forward in recognizing its vision. The DON chief data officer asserts that the DON's information eventually will be consolidated into a greater Navy enterprise data lake, where data "ownership" is not emphasized, but rather where high quality, usable data will be made available to relevant parties [11]. In this environment, machine learning and AI algorithms can be applied to this data in order to make better data-driven decisions, therefore moving the DON closer to its strategic goals regarding data management.

All of this leads to Jupiter, which is the DON enterprise data environment launched in April 2020. This platform makes data available and usable across the naval enterprise, and it provides advanced data tools, context-rich visualizations, and decision-support analytics [12]. Jupiter is "at the heart of the DON Data Architecture," and it is the realization of the DON's data infrastructure modernization efforts [12]. The DON chief data officer is currently looking for new naval use cases for this platform, and COVID Fleet Readiness is one such use that has been proposed [11], [12].

2. COVID-19 Fleet Readiness using Jupiter

In an effort to consolidate reporting and analytical efforts, the Navy is developing a way to leverage this modern DON data architecture in order to consolidate COVID reporting to the Office of the Chief of Naval Operations (OPNAV), standardize COVID reporting formats, fuse data, and integrate AI/ML analytics into decision-making. The

system is meant to incorporate COVID testing data, supply and logistics data, and personnel data into the Jupiter platform as seen in Figure 1 [13].

Figure 1. Data Flow into Jupiter from Various Databases in the DON. Source: [13].

3. Integration of Contact Tracing Data into Jupiter

There is no definitive source that claims that the collected proximity data will directly be used in Jupiter as part of a larger fleet readiness system, but it is reasonable to assume that if the DON is to incorporate COVID-19 incident and test data into this system, the actual contact tracing data of individuals could easily be incorporated as well. This follows from the DON's modernization efforts discussed earlier. For the purposes of this thesis, we assume that contact tracing proximity data will be collected using a DON procured system and that the collected data will eventually be integrated into the Jupiter platform.

The model for the system architecture referred to in this thesis and pictured in Figure 2 will include the wearable devices and the backend Jupiter platform that processes

the data. This will serve as the model architecture that will be evaluated using the NIST Privacy Framework.

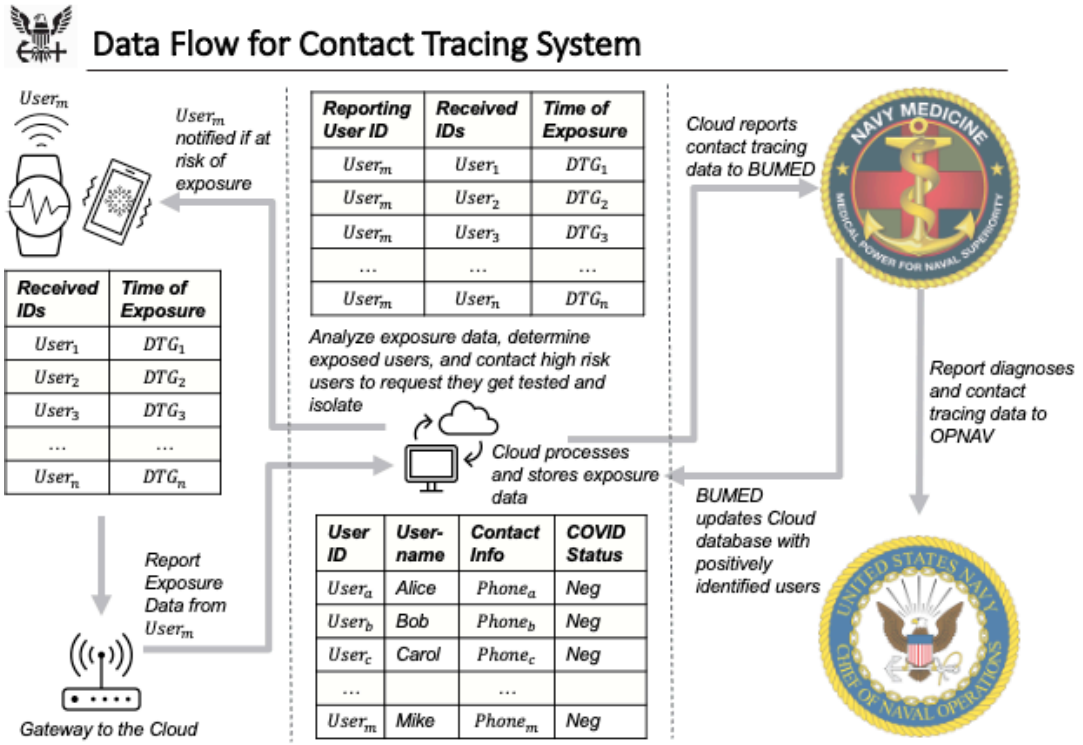

Figure 2. Data Flow for the Collected Contact Tracing Data

C. NIST PRIVACY FRAMEWORK

This section gives an introduction to the NIST Privacy Framework that will be used to evaluate the DON contact tracing system. The NIST Privacy Framework is the result of numerous laws, reports, and federal publications. The timeline leading up to the framework reflects the U.S. government's focus on preserving digital privacy in the same way it focuses on information security, and it also shows the effort that is being expended toward protecting citizens' privacy in the wake of emerging technologies that pose a threat to privacy.

1. From the Privacy Act of 1974 to the NIST Privacy Framework

On January 16, 2020, the National Institute of Standards and Technology (NIST) released "NIST Privacy Framework: A Tool for Improving Privacy Through Enterprise

Risk Management, Version 1.0," which we refer to as the NIST Privacy Framework [1]. The release of this framework is the culmination of years of efforts to move privacy to the forefront of the national conversation. Beginning with the Privacy Act of 1974, the United States has long had deliberate privacy regulations in place, but in recent years with the exponential increase in available data and advances in technology to analyze this data, the United States has had to continually reconsider and update its approach to privacy and how privacy regulations reflect the country's commitment to the privacy of its citizens.

Emerging trends in the technological landscape such as big data in government, large scale collection of data by third parties, and metadata versus personal data have prompted the government to solidify their stance on digital privacy. In May 2014, the White House released two documents titled "Big Data: Seizing Opportunities, Preserving Values" and "Big Data and Privacy: A Technological Perspective" [14], [15]. Both documents reinforced the need for privacy enhancing technologies in the era of massive data collection, and they both provided recommendations for writing policy and regulations that reconcile the government's stance on privacy with the privacy threats that come with Big Data. Both of these reports shape the context for a DON contact tracing system as the volume and the velocity of the data collected and processed has the ability to infringe on user privacy.

The same year those reports were released by the President, Congress enacted the Federal Information Security Management Act of 2014 (FISMA), an update to the Federal Information Security Management Act of 2002. FISMA 2014 is primarily focused on information security, "provid [ing] for development and maintenance of minimum controls required to protect federal information and information systems" [16]. This is important because it shows that the U.S. government is taking an active role in administering the implementation of information security policies.

In addition to the minimum controls requirement required by the law, FISMA requires the Office of Management and Budget (OMB) to update its Circular No. A-130. In 2016, the OMB released a revision of OMB Circular No. A-130 to reflect changes in laws and advances in technology. The revision "establishes general policy for planning, budgeting, governance, acquisition, and management of federal information, personnel,

equipment, funds, IT resources, and supporting infrastructure and services," all while placing an emphasis on security and privacy in the federal information life cycle [17]. These policy directives are further implemented in "Fiscal Year 2019–2020 Guidance on Federal Information Security and Privacy Management Requirements," released in November 2019 as a Memorandum for the Heads of Executive Departments and Agencies. The memorandum highlights the importance of complying with privacy requirements and managing privacy risks. It also provides guidance on agencies' privacy programs and requirements for agencies to report the status of their information security programs to the Office of Management and Budget [18].

Building on the OBM Circular A-130 emphasis on managing privacy risk, the National Institute of Standards and Technology (NIST) issued NIST Internal Report (NISTIR) 8026 "An Introduction to Privacy Engineering and Risk Management in Federal Systems" [3]. This document further describes privacy engineering concepts that need to be developed and implemented in federal system in order to fulfil the requirements of OMB Circular A-130. Ultimately, this document points toward expanding guidance and integrating privacy engineering into existing NIST Risk Management Frameworks (RMF).

One of these RMF documents, NIST Special Publication 800-37 rev2 "Risk Management Framework for Information Systems and Organizations: A System Life Cycle Approach for Security and Privacy," issued in December 2018, includes updates to "integrate privacy risk management processes into the RMF to better support the privacy protection needs for which privacy programs are responsible" [19]. Additionally, NIST Special Publication 800-53 rev5 "Security and Privacy Controls for Information Systems and Organizations" released in September 2020 continues to integrate information security and privacy into a detailed list of controls, policies, and procedures [6]. While both of these publications provide controls and recommendations for security and privacy, NIST determined that organizations need frameworks to guide their use of these publications and streamline the control selection process, and this leads to the release of the cybersecurity and privacy frameworks.

The NIST publications mentioned above are recommended for ensuring federal information systems comply with U.S. government security and privacy requirements, but

due to government's push for enhanced cybersecurity and privacy in non-government sectors as well, NIST has been entrusted with developing more flexible cybersecurity and privacy risk frameworks for voluntary use by both public and private organizations. This started with NIST releasing the "Framework for Improving Critical Infrastructure Cybersecurity v1.0" in February 2014 (with a rerelease of v1.1 in April 2018) as the result of the Cybersecurity Enhancement Act of 2014 (CEA) and Executive Order (EO) 13636, "Improving Critical Infrastructure" [20].

This act and Executive Order were the basis for NIST refining its role in information security by developing a cybersecurity risk framework "for voluntary use by owners and operators of critical infrastructure" [20]. This framework "offers a flexible way to address cybersecurity" and the effects it has on organizations [20]. The NIST Cybersecurity Framework does address privacy protections insofar as violations to privacy may result from cybersecurity flaws, but the framework also acknowledges the significant difference as well as the connection between cybersecurity and privacy. This was a large driver in NIST following up the cybersecurity framework with NIST Privacy Framework [1]. Figure 3 shows a timeline of the major publications that have led to the NIST Privacy Framework.

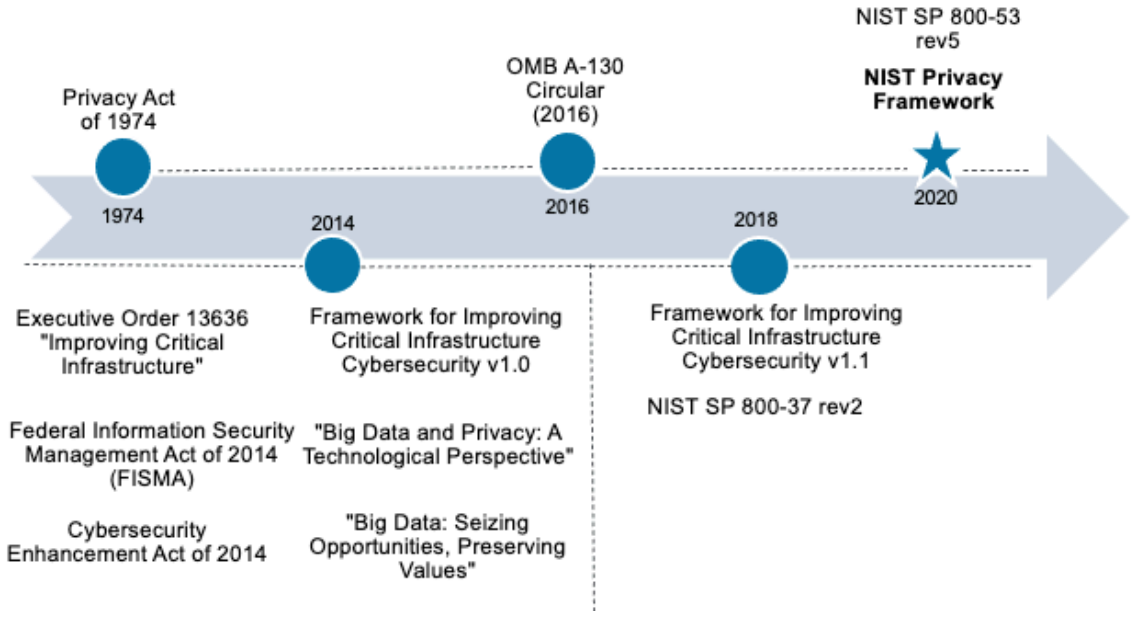

Figure 3. Major Publications related to the NIST Privacy Framework

2. NIST Privacy Framework

The NIST Privacy Framework is similar in structure to the Cybersecurity Framework, and this is to facilitate organizations being able to use them in tandem. The privacy framework acknowledges the challenging nature of defining privacy and implementing privacy controls that might not fall purely within the realm of cybersecurity; therefore, the framework is intended to be usable in a wide variety of organizations regardless of size or technological focus.

Like the Cybersecurity Framework, the Privacy Framework consists of Core, Profile and Implementation Tiers. The Core lists describes privacy activities as categories and subcategories. NIST has also provided a mapping, separate from the framework document, of these categories to controls in NIST SP 800-53 rev5. These Core activities are meant to open a dialogue within an organization about privacy risks and what areas and levels of risk the organization is willing to assume. The Profiles provide a way to construct organization's current risk profile based on the Core activities previously identified. Then the organization can also construct a target profile showing where they want to be in terms of a privacy assessment. Finally, Implementation Tiers are used to evaluate how well an organization is adequately managing its privacy risk [1]. Figure 4 gives the NIST definitions of the Core Profiles and Implementation Tiers.

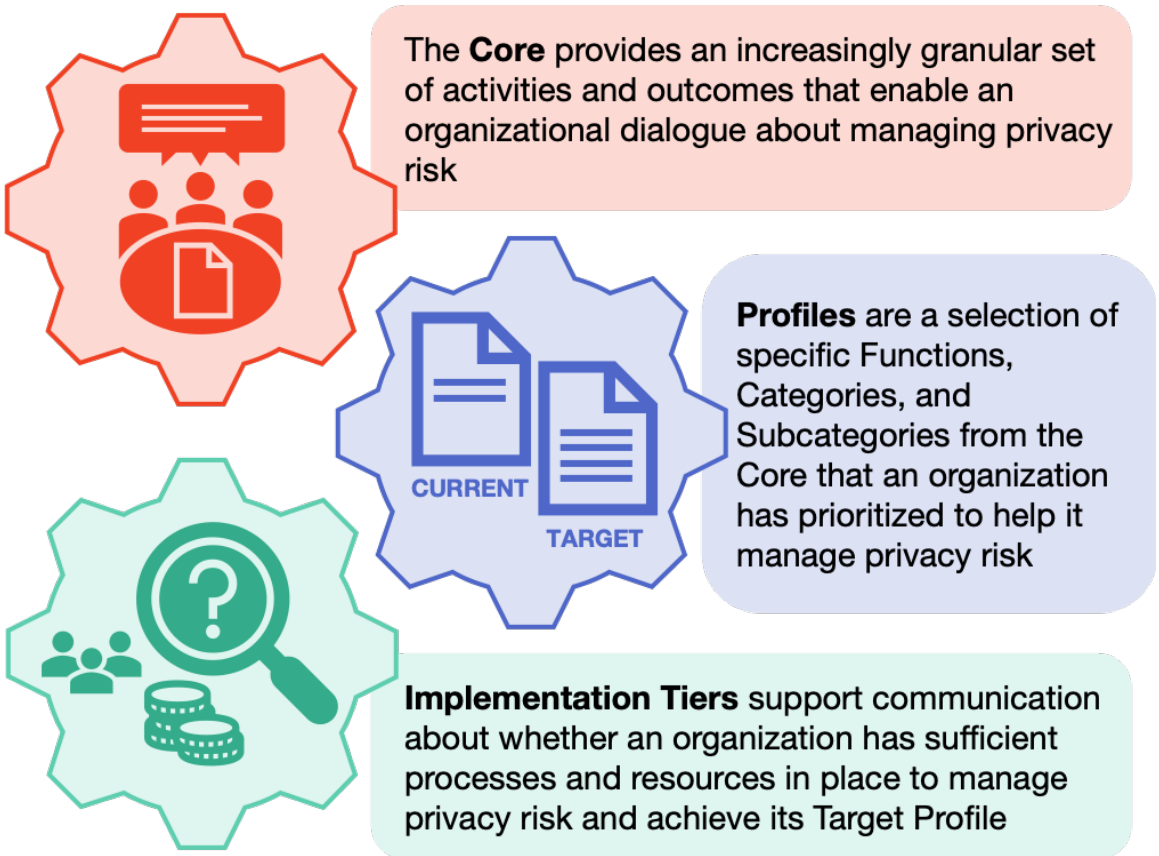

Figure 4. NIST Privacy Framework Structure. Source: [1].

The NIST Privacy Framework is a starting point for organizations to make a concerted effort to consider and manage privacy risks. Starting this dialogue and talking about privacy risks in an organization should lead to implementing proper controls. This is where the NIST 800 series special publications can be integrated into the ongoing process of preserving privacy in an organization.

The DOD requires that cybersecurity requirements follow an RMF that is consistent with NIST SP 800-37 according to Department of Defense Instruction 8510.01 [21]. NIST SP 800-37 references the Cybersecurity Framework in its controls, but not the Privacy Framework. The Privacy Framework is relatively new, and it is imperative that it be applied and evaluated thoroughly before being added to NIST Special Publications in the same manner as the Cybersecurity Framework. NIST anticipates issues and states the "wealth of resources does not yet exist in the privacy domain" as it does with cybersecurity, and organizations normally "find it challenging to integrate privacy risk assessment into their

risk-management approach" [22]. This is why it is so important that this framework eventually be molded into something that provides the utmost utility to organizations in applying privacy controls .

III. SYSTEM PRIVACY CONSIDERATIONS

This section explores the different privacy considerations of a DON contact tracing system starting with a description of the DOD organizations that are responsible for privacy. It is important to establish a baseline of what privacy resources and mechanisms are already in place at the federal level because this will affect how the NIST Privacy Framework will be able to integrate into an already established privacy program. This also highlights the organizations that would need to be involved, either directly or indirectly, in ensuring privacy in a contact tracing system. The breadth of organizations involved points to the scope and scale of the system where data from the system may pass through multiple departments and agencies.

While the granular details of the contact tracing system are not discussed in this paper due to the focus on its overall privacy concerns and how the NIST Privacy Framework can integrate into a system of this scope and scale, this section describes the prominent privacy considerations inherent to the different parts of the system. This includes privacy concerns with collection by the wearable devices, and it also includes privacy concerns inherent to how the data is stored and processed after collection.

A. DEPARTMENT OF DEFENSE PRIVACY

This section details the organizations and some of the resources involved in privacy assurance for the DOD. It is necessary to discuss these entities since their function contributes to our understanding of how the contact tracing system and NIST Privacy Framework would interact with established privacy offices and programs.

1. Defense Privacy and Civil Liberties Programs

At the highest level within the Department of Defense, privacy is overseen by the Defense Privacy & Civil Liberties Division which administers the DOD Privacy and Civil Liberties Programs in accordance with DOD Instruction 5400.11 [23]. This instruction details the roles and responsibilities for the DOD Privacy and Civil Liberties Program, but also addresses key privacy principles when collecting personally identifiable information (PII). This includes directing DOD components to limit the “creation, collection, use,

processing, storage, maintenance, dissemination, and disclosure of personally identifiable information (PII)" to that which is "necessary to accomplish a DOD function," all of which come into play for contact tracing systems [23]. Additionally, this instruction requires DOD components to impose conditions when sharing PII with other federal or non-federal agencies as well as entities "that govern the creation, collection, use, processing, storage, maintenance, dissemination, disclosure, and disposal of the PII" [23]. Furthermore, the data collected by the Navy's contact tracing system will likely be transferred to different owners that may or may not fall under the authority of the Defense Privacy & Civil Liberties Division.

The most important contribution of the Defense Privacy & Civil Liberties Division is the policy stated in section 1.2.a.(1) of [23] that requires all DOD components to:

> Establish and maintain comprehensive privacy and civil liberties programs that comply with applicable statutory, regulatory, and policy requirements, and develop and evaluate privacy and civil liberties policies and mange privacy risks.

This is significant because having a privacy program unique to each DOD component allows for more detailed regulations and control of the privacy measures of the systems owned by those components; however, this can complicate matters in the event that PII data moves among systems administered by different components or shared by several components.

2. Privacy Impact Assessments

Regardless of which component of the DOD that PII resides with, every federal agency is required by the E-Government Act of 2002 to complete a Privacy Impact Assessment (PIA) for procuring technology that "collects, maintains, or disseminates information" that has PII implications [24]. All PIAs are publicly available unless their classification or sensitivity does not allow for public release. The purpose of PIAs are to show that privacy considerations have been taken into account throughout the entire life cycle of the system [24].

PIAs address questions relating to PII collected by a system mainly from a legal and accounting perspective. It is a standard form that include sections describing the type

of data collected, the authority to collect the data, the individuals' autonomy surrounding their data, and whether the proper privacy notices have been provided to individuals. Relevant DOD components, agencies, and contractors with access to the system's data are listed on the form along with information on how they will interact with the data in the system.

While PIAs cover many important areas required to assess the privacy of a system, they do not cover many technical specifications of the system that may cause underlying privacy concerns. That is to be expected since these forms appear to be more for compliance purposes while the technical privacy considerations are left to the NIST Special Publications that detail the privacy controls that should be used in federal systems. The NIST Privacy Framework should be able to merge these two areas—consolidating the technical and legal underpinnings of privacy controls in federal systems into a holistic risk-management view. Currently, PIAs appear to be the DOD requirement that is most similar to the NIST Privacy Framework in reference to it being a privacy assessment of a federal system; however, the NIST Privacy Framework provides a more detailed roadmap for evaluating the privacy of a system.

3. Department of the Navy Privacy Program

The Department of the Navy Privacy Program is integral to the successful development, operation, and sustainment of the Navy's contact tracing system because the data collected will likely be processed and analyzed by the DON's data management system, Jupiter. In accordance with DOD Instruction 5400.11 mentioned above, the Department of the Navy Privacy Program was established by SECNAV Instruction 5211.5F [23], [25]. The DON Privacy Program mirrors some of the broader policies of the DOD Privacy Program, but the DON includes more specific directives when dealing with PII.

It is important to note that SECNAV Instruction 5211.5F directs the DON to protect information from unauthorized access or disclosure, and the DON must safeguard information with the proper technical controls [25]. Although this heavily alludes to security controls, this clearly affects privacy, and this is the sort of thing that the NIST Privacy Framework was designed for. It was meant to assess and determine controls for

situations where security and privacy overlap [1]. When using the NIST Privacy Framework to evaluate this contact tracing system, the system's stakeholders need to address these security and other controls from a privacy perspective, and this is more nuanced than just adding encryption to all data collected by the system or ensuring authentication of all users that access collected or analyzed data.

4. DHA and BUMED Privacy Programs

Another agency that is likely to be a part of the DON Contact tracing system is the Defense Health Agency (DHA). We assume that the contact tracing data may be sent to a DHA subordinate command like the Bureau of Medicine and Surgery (BUMED), the health care agency for the DON, or that the contact tracing data could be merged with medical data from BUMED. When a personal health information (PHI) is collected, stored, or processed, it falls under the authority of the Health Insurance Portability and Accountability Act (HIPAA), and this includes its own set of privacy standards that are continuously updated to keep up with evolving technology [26]. This is the most significant factor that separates the DHA Privacy and Civil Liberties Office from the other DOD components, and it needs to be taken into consideration when applying the NIST Privacy Framework to the DON contact tracing system which will likely process PII and PHI.

The DHA Privacy and Civil Liberties Office has a detailed and well documented privacy program. The DHA Privacy Program Plan lists the federal laws and DOD regulations that the DHA is subject to. This includes NIST SP 800-53 "Security and Privacy Controls for Federal Information Systems and Organizations," the RMF guide required by DODI 8510.01 "RMF for DOD IT" [26]. The Privacy Program Plan even provides a Privacy Controls Mapping to CNSS 1253 (an instruction that builds on NIST SP 800-53), and it also provides a Table of Requirements to aid compliance to the DHA privacy program authoritative documents [26], [27].

The NIST Privacy Framework, if applied properly, should be able to reconcile the differences, if any, in the privacy programs between the DON and the DHA. Although both programs adhere to the same federal laws and regulations, with the exception of DHA adhering to additional PHA privacy regulations, the data flowing between systems

administered by these two entities needs to be subject to privacy controls that are commensurate with the type and sensitivity level of the data.

B. PRIVACY TRADEOFFS IN THE DON CONTACT TRACING SYSTEM

Since the outbreak of COVID-19, there has been a concerted effort within and outside the healthcare community to obtain various types of data, including personally identifiable information, that can be used to develop and implement measures to defeat the COVID-19 pandemic (for a high-level description of the most prominent digital contact tracing systems in use, refer to Appendix A). Digital contact tracing is viewed by the healthcare industry and governments as an important addition to manual contact tracing. The thinking is that automation of the collection and analysis of COVID-19 data will reduce the time from sensing or predicting outbreaks to when governments and individuals can take action to avoid otherwise severe consequences. However, there are ethical considerations surrounding the use of automated contact tracing systems, particularly with regard to the potential for misuse of the collected data and the collectors (i.e., recipients) of the data not being fully transparent the people on whom data is being collected [28].

Most of the discussion around ethical dilemmas of operating digital contact tracing systems is centered on privacy. Researchers have delved deeply into this subject by analyzing the privacy concerns of some of the more prominent digital contact tracing systems used around the world. For a detailed description of these concerns, refer to Vaudenay (2020) [29], [30]. Privacy concerns about contact tracing system have impeded the speed at which they have been adopted; therefore, these systems cannot fulfill the claimed purpose of stopping the spread of the virus. Some of the contact tracing systems detailed in Appendix A like the Apple/Google system have tried to use privacy-protective features as an enabler to facilitate adoption of their system, but privacy is still a large concern [31]. MIT launched a tracker in May 2020 which looks at the policies and safeguards surrounding different countries' digital contact tracing systems [32]. In a broad sense, this shapes our assessment of the DON system as well since it highlights some attributes of contact tracing systems that are important to the public like their voluntary nature and their transparency, and presumably the DON system will be mandatory if it is officially rolled out in the fleet.

The DON has already outlined some of the specifications that it is requesting in the design of its contact tracing system to include a wearable device and the use of Bluetooth over GPS for proximity measurements [7]. The following sections cover some of the privacy tradeoffs afforded by system design choices that either the DON has already made or might consider during later acquisition phases. Privacy affordances can result from design choices or be inherently linked to the nature of digital contact tracing systems. The privacy effects of such a system may create unwanted outcomes that affect the ability of a DON command to conduct its mission. To the extent that design choices mitigate these unwanted outcomes, we examine the tradeoffs below.

1. Centralized versus Decentralized

The main privacy tradeoff for contact tracing systems is whether the system is centralized or decentralized. The distinction between centralized and decentralized systems is vital to analyzing the privacy preserving nature of the contact tracing systems in development and use. Centralized and decentralized systems can best be explained in terms of who has power over the data or how the data is controlled and used.

Centralized contact tracing systems generally involve having a primary data center operated by public or private healthcare authorities. The data center serves as the center of gravity for data storage and processing, and the authority operating the data center has access the ability to process the incoming data and potentially re-identify data that is being collected. Consider the following high-level description of an example of a tracing system founded on a centralized computing scheme. When a user is registered with the system, the data center server issues the user a cryptographically generated temporary identifier that will be broadcasted via Bluetooth to other users within Bluetooth range [33]. If a user is diagnosed with COVID-19, and the user uploads the temporary identifiers and associated metadata (e.g., time and estimated distance) observed from other devices to the central data location, authorities that control that data repository are able to determine how that proximity data is processed. They also control the ability to determine which users are now at risk and notify those users to take appropriate follow-on actions (medical testing, self-isolation, etc.) [34]. All the while, the authority overseeing this system has assess to user data flowing through the system.

Decentralized systems, on the other hand, give the users more agency over their own data. In a completely decentralized system, there exists no central authority that has the power to utilize the data being collected and stored throughout the system. Essentially, end users have access to only their data (usually stored in their end devices), and they do not have control over or access to the data of others.

In decentralized contact tracing systems, the temporary identifiers that users broadcast are usually generated cryptographically in the users' devices derived from a seed known only to each device. If a user is diagnosed with COVID-19, they can upload the seed used to generate their temporary identifiers to a backend server. Other devices periodically query this backend server for seeds from positively diagnosed users. When a device downloads those seeds, it can generate the temporary identifiers locally and compare them to temporary identifiers that they have stored locally in their devices over the relevant exposure period. If their device finds a match between the temporary identifiers computed from received seeds and the temporary identifiers observed, the device can establish that the user might be at risk due to exposure, and alert the user to take appropriate measures [30]. Yet the backend server, which knows only the seeds reported by individuals who have opted to disclose their positive test status, cannot establish when or where in the network of proximity relationships between users any specific interactions took place. Further, the server has learned no information from users who have not disclosed a seed, and its capacity to make secondary use of proximity data is thereby limited.

The aspect of centralized and decentralized contact tracing systems, as it relates to this assessment of the DON system, revolves around the control and the ability to use the collected data for decision making purposes. The description of the DON system in the RFI fits squarely in the realm of centralized systems. The DON even states that it wants to use the collected data to determine if social distancing policies have been working [7]. The benefit to the DON using a centralized contact tracing system lies mainly with the ability of the DON to control access to and use of the data. A centralized system can provide more data to healthcare workers an decision makers if it is possible to observe a more complete representation of the social graph between users than contact tracers might otherwise have [30]. Centralized systems can also allow for a human in the loop (discussed in Section

III.B.3) to augment the analysis and evaluation of the spread of the infection [33]. Another advantage of centralized systems is that it enables DON workers to ensure the data is available and actionable to the parties that need it and to control secondary uses of data in explicit ways. Otherwise, data could remain siloed in separate collections, potentially countering analytical benefits of collecting proximity data in the first place.

The privacy tradeoff with the DON using a centralized system revolves around the DON's ability to potentially make secondary use of data provided (for example, inferring location from proximity information or using proximity data to investigate conformance with social distancing requirements). A central server could also be breached by an adversary (or insider threat), disclosing data to unscrupulous entities; such breaches could even introduce a backdoor mechanism that converts temporary identifiers to associated long term identifiers or otherwise deanonymize users. While the ability to pull off an attack that could compromise the server and backdoor might be difficult to achieve, the impact of a successful attack would be severe [30].

2. Wearable Device versus Phone App

Another tradeoff concerning the privacy of the DON contact tracing system is how wearable devices and phone apps interact with the system. One of the most prominent privacy concerns with phone apps is the amount of private data as well as the number of sensors present on all smartphones. Although the contact tracing apps described above claim to protect private data by only broadcasting pseudorandom temporary identifiers, it would not be too crazy of a thought that if a poorly developed app were somehow compromised, it could exfil sensitive data located on the user's device [35].

In contrast to a smartphone app, a wearable device's capacity for privacy is based on the amount and type of the data stored on the wearable device. No personal data should be stored on the device, and the device should not maintain a persistent internet connection. This means that there would be no risk of extraneous data being transmitted by the wearer [35]. The only data capable of being transmitted would temporary identifiers and their associated metadata (signal strength, contact time, duration, etc.) used by the system for making risk assessments.

There may have been several reasons why the DON has chosen to procure a wearable device rather than develop or use a phone app. These could include cost, simplicity, or more control over the system. Either way, privacy assurance of using either a wearable device or a phone app is based largely on privacy controls that ensure that only data necessary for the operation of the system is able to be collected and that users are aware of what data is being collected from them. The DON needs to consider these controls and their impacts when ensuring what type of data is collected during operation of the system.

3. Human in the Loop versus Automated

With any contact tracing system, there is the concern of the extent of human interaction in the system versus automation. There is obviously a benefit to having heavy human interaction in the system. The system implemented by Singapore (wearable and smartphone app) was chosen over the Apple/Google system because it gives Singapore health officials the ability to keep humans in the loop of contact tracing [33], [36]. The developers of their system acknowledge the privacy benefits of using an automated notification system, but they also understand the need for human contact tracers to be included in the system to be able to incorporate other outside information besides just physical proximity data [33]. The human contact tracers can incorporate data that is not collected by the app such as locations and environmental conditions.

Human-in-the-loop and automation approaches when designing a system like the DON system are not a "one or the other" decision. It is a spectrum approach, and each organization has the ability to determine how much of the system they want automated and how much human interaction they require. There are essentially two areas that where a human in the loop can play a major role. Humans can be used for exposure notification and treatment and humans can be involved in secondary uses of data to produce aggregated readiness products.

Since the DON wants to use COVID-19 data for decision-making and fleet readiness purposes, it makes sense that they would want to incorporate contact tracing data into the decision-making process. A more human-in-the-loop approach allows these decision makers to have more of a say in this process. This contact tracing data can be

evaluated by decision makers and merged with other COVID-19 related datasets to include manual contact tracing data. This will result in cleaner and more complete data that can then be fed into the Jupiter data management system. Once again, this increased human involvement and access to collected data has the potential to raise privacy concerns, but that privacy risk has to be weighed in light of the data's value and benefit to naval operations, and the DON has to try to mitigate these privacy concerns as much as possible.

4. Location Based versus Proximity Based

Proximity based methods are the more privacy preserving choice for a contact tracing system. In a world where privacy does not matter, location-based methods would certainly be the preferred choice in most environments as it would be able to tie location data to proximity data. Countries like China, Israel, and South Korea use citizen's electronic location data for contact tracing; therefore, some countries believe the privacy tradeoff is insignificant compared to the cost of the pandemic lockdown [37].

There is still a debate about whether location data is even more beneficial, necessary, or even useful in contact tracing since the only data that seems to matters is whether people are in close enough contact to for a period time to spread the infection [37]. Regardless, none of the most prominent digital contact tracing systems include location-based data in their design, and this consideration is likely due to privacy concerns and concerns about the willingness of users to support adoption of a system if it tracks their location data—data which could be viewed by governments, health authorities, or other users.

The DON chose the more privacy preserving option, but the Bluetooth tracking systems are not without privacy concerns. Back in 2004, MIT researchers used the Bluetooth capability of mobile phones as "wearable sensors" along with machine learning techniques to determine information about users' patterns of activity. They were able to reconstruct user daily movement patters as well as infer friendship relationships between users [38]. The MIT paper shows that proximity data has significant value, and this value increases when linked to ancillary data. This poses an inherent threat to collecting proximity data from DON employees, and this plays into the potential privacy pitfalls of

the DON contact tracing system that need to be accounted for in a privacy risk assessment and mitigated by privacy controls.

C. FLEET READINESS SYSTEM PRIVACY CONSIDERATIONS

Once proximity data has been collected from the wearable contact tracing system, the data may be stored locally at each command, but its utility to the Navy is mainly predicated on its ability to be integrated into the larger DON enterprise. This integration will allow that data to be processed and analyzed using advanced analytic techniques, but it also allows the sharing of the data between entities. Making this data available for sharing between DON entities will enhance the analytic capabilities of the system, and it will provide more detailed and accurate fleet readiness products to high-level decision makers in the DON.

This backend system that stores and processes the data comes with inherent privacy risks that need to be addressed. Most of these risks are due to the scope and scale of this system and would be inherent to any datasets that are absorbed by the DON's enterprise management system. This section will highlight the larger privacy issues facing the use of this system to process contact tracing data. Identifying the privacy concerns of this system is necessary to applying the NIST Privacy Framework, which is done in Chapter IV.

1. Data Governance by Multiple Organizations

In a system of this scope and size, the data collected is going to traverse several subsystems that belong to different organizations, which means that the overall system will fall under several different, usually overlapping, privacy programs. This aspect of data governance affects the buying decisions of the DON since the DON needs to ensure contractors involved in this system adhere to DOD privacy regulation. The NIST Privacy Framework addresses this as the framework attempts to give guidance on how to deal with situations where the purchasing agency has to choose between several different suppliers that have different sets of privacy requirements or their privacy requirements are different than the purchasing agency [1].

This concern is applicable when initially purchasing the wearable contact tracing devices and storage devices. The NIST Privacy Framework can help the DON determine

their list privacy requirements, and this can be used as a template for selecting a vendor for the devices. Any privacy gaps determined to be in the vendor's system will need to be weighed against the privacy risk willing to be assumed by the DON. The DON already does this for security when procuring new systems. In August 2020, Naval Surface Warfare Center Crane Division assisted with providing security assessments of a potential contact tracing system that was to be procured by the DON for trials at the United States Naval Academy (USNA). While security assessments of a system are important from a privacy perspective as well, attention should also be given to aspects of the system that pose a privacy threat in addition to the security of the system.

It is important to note that the DON Privacy Program does address contractor privacy responsibilities, and the idea is that they comply with federal privacy regulations and receive the proper privacy training [39]. This sets the foundation for legal compliance for contractors when handling PII and PHI, but there should be some other additional mechanism that aids contractors and the DON in ensuring privacy in systems rather than just putting a check in the box that the system is legally compliant. There appears to be no technical guidance about data governance by Contractors, and this makes it difficult for the DON to guarantee privacy of data across multiple systems that involve different contractor software and hardware.

One of the main considerations around privacy and data governance that will arise from this system's architecture will be the transference of data to and from the Military Health System (MHS) and the DON's subcomponent, BUMED. The DON is already proposing a system where BUMED data is integrated into the Jupiter data management system in order to provide fleet readiness information to the Office of the Chief of Naval Operations (OPNAV) [13]. If contact tracing data is to be sent to BUMED from each command before being integrated into the Jupiter system, that data will go from being PII to PHI since DHA oversees "covered entities" under HIPAA which include health care providers that can handle health care data in electronic form [26].

When the contact tracing data is collected at a Navy command, the data must be treated as PII if it contains pieces of information that are unique identifiers for employees. After contact tracing data is generated from the wearable devices, the contact tracing data

has to be linked to a personal identifier at some point in the system. This is so that the wearer of a device can be contacted if they are at risk of having contracted COVID-19. If this contact tracing data is sent to BUMED, it can be linked with PHI data about the individuals. This gives health professional more information to work with in analyzing the spread of COVID-19. This also gives OPNAV via the Jupiter system more data to work with in making predictions about fleet readiness; however, when the contact tracing data is released from BUMED, the data is no longer covered under HIPAA.

Essentially the DON needs to be able to use the NIST Privacy Framework to reconcile multiple subsystems of the larger contact tracing system in order to ensure that technologically their privacy controls are on the same level. Many of the same legal and regulatory statues apply to organizations working with or within the DOD; therefore, a focus should be placed on technological controls as subsystems interact. The NIST Privacy Framework needs categories to identify these potential privacy controls gaps and be able to lead system engineers in a direction that will lead to adequate technical privacy controls.

2. Distributed Architecture

This section highlights the concerns with the contact tracing system being spread out across the DON enterprise. If contact tracing data was to remain local to each command, these concerns would not be as worrisome; however, the RFI for the system did request some type of class based storage architecture [7]. This implies that a more distributed nature, and privacy controls need to be implemented accordingly.

a. Cloud Architecture

The contact tracing system will use a cloud architecture for storing processing, and transferring data, as the DON is planning on using Microsoft Azure services as part of the system [13]. In the RFI for the wearable contact tracing system, it also mentions the possibility of the collected data being stored in a cloud-based architecture [5]. Security and Privacy have been studied extensively in in cloud computing in recent years [40], [41]. While the research overwhelmingly discusses cloud security, cloud privacy receives less attention, and is often cited as ancillary to the architecture but necessary to ensure trust in

the overall system or application. The following paragraphs represent some of the privacy concerns related to the cloud storage aspect of this system.

While access control usually falls under security measures, it also fits squarely in the realm of privacy measures, and this is particularly applicable to a large cloud-based system that will likely have different levels of privacy requirements. This can be a complicated feature of a large system like the one being described in this paper. The contact tracing data stored in the cloud can be integrated with other types of data that may not have the same privacy requirements, and this could further complicate access control, and access control lists can become unmanageable [42]. This issue of access control is fundamental to any system, and it is important to see if the NIST Privacy Framework can walk developers of this contact tracing system toward a technically feasible, economically viable, and fit-for-purpose access control solution that best preserves privacy.

Another privacy aspect of accessing data in a cloud architecture is the ability of an adversary to infer user behavior when accessing sensitive data [40]. While unauthorized users accessing sensitive contact tracing or medical data can directly lead to OPSEC concerns, this vulnerability of inferring user behavior is easier to overlook. Information can be leaked indirectly by observing the pattern or timing in which someone else accesses the data. This could lead to an adversary or unauthorized user gathering information about fleet readiness or other operations.

Using a cloud-based architecture also increases the difficulty of conducting audits on the system [42]. Audits ensure privacy policy compliance but conducting audits may be difficult on a system of this nature since there may be multiple privacy policies based on the type of data stored (PII versus PHI). Auditing would have to scale to several subsystems, and this would raise the question of whether or not the audit mechanism can be integrated at the command level where the data is collected and can also be integrated within the Jupiter and Azure architectures. The other option would be separate custom audit mechanisms needed in each subsystem of the architecture.

b. Priority of Data

The distributed nature of this system leads to questions of whether or not the NIST Privacy Framework is effective in helping the DOD to conduct an adequate risk assessment when different nodes in the architecture may have different risk tolerances. The stance of the DON CIO must be to eliminate unauthorized disclosures of personal information whenever possible, but some personal data is more critical to operations than other personal data, and this should play into the risk calculations when assessing the privacy risk of this contact tracing system.

For example, the readiness of a Carrier Strike Group (CSG) that is about to deploy is likely to be a higher priority than the readiness of the students at the Naval Postgraduate School (NPS) or the United States Naval Academy (USNA), but ultimately it comes down to a discussion of risk. It is easy to automatically think that the leaking of a COVID-19 outbreak in a CSG would constitute more risk to operations, but that is not completely true. While the priority of the operations of a CSG might be higher than that of operations at the USNA or NPS, likelihood and impact also play a role in determining risk. Priority of data is just a single, but important factor in determining risk. This priority of data combined with the likelihood and impact of inadvertent disclosure could tip the scale and cause leadership to increase or lower the privacy risk tolerance of that data (this is discussed deeper in Chapter IV). If the DON determines that certain data is more valuable to fleet readiness and operations, then the DON will have to accept more privacy risk regarding this data, as it is more likely to reveal valuable readiness information if disclosed.

The same way that the NIST Privacy Framework should lead users to being able to make a risk assessment of a system involving multiple entities (in this case, government agencies and contractors) that may interact with the system, it should also be able to assess the criticality of data being collected so that it can be prioritized when calculating privacy risk tolerance.

3. De-identification

De-identification is one of the quintessential focus areas of privacy research much in the same way that encryption techniques are the focus of a lot of security research. De-

identification is a process where data is modified so that individuals cannot be specifically re-identified and others cannot gleam information about an individual using other attributes about him or her [43]. Research has shown that the capability exists to re-identify large data sets. One prominent example would be the de-identification of the Netflix Prize dataset, where researchers were able to identify individual subscribers from a set of 500,000 subscribers on Netflix. They were able to correlate this data the Internet Movie Database (IMDB) data to even reveal other potentially sensitive information about users [44]. Another set of researchers were able to identify 95% of individuals using only four spatio-temporal points from their cellphone data [45]. These examples are meant to show that de-identification has become a major concern surrounding big data sets and individual privacy, and since the contact tracing system will feed the contact tracing data into AI/ML algorithms for processing, the DON needs to take the proper precautions to limit re-identification if data has been de-identified.

One of the issues surrounding de-identification that needs to be addressed is the tradeoff between the privacy of the data and the utility of the data. The DON will eventually have to make such a tradeoff because guaranteeing a certain level of privacy in terms of data e-identification can lead to a decrease in the utility of the data being processed [42]. The DON needs to clearly define an upper bound to privacy controls and that needs to be based on the utility that they expect from the data being processed. In other words, if too much identifying information is removed from the data collected the data will be of limited use in some forms of predictive analysis. The DON needs a framework or some kind of metric to analyze the privacy versus utility tradeoff in this contact tracing system.

Since the contact tracing system is likely to be centralized system (as described in section III.B.1.), this brings with it the tradeoff of centralized versus decentralized de-identification [42]. Centralized de-identification may be a more complex approach for the DON as the DON is given the authority to de-identify data as it sees fit as long as de-identification is done in accordance with regulations. The DON would then need to make the tradeoff described earlier about how much de-identification will affect the use of the data. On the other hand, if a decentralized architecture were proposed for the wearable devices, the contact tracing data could be de-identified locally, but this would limit the use of the data as now identifiers have been removed that could have been used to link other

data sources (i.e., BUMED data). It makes sense that the DON would want to use a centralized system, but there needs to be a mechanism or a requirement on how they anonymize the contact tracing data. The NIST Privacy Framework should address at least some aspect of this.

The final major consideration about de-identification revolves around the challenge of streaming data [42]. With the rise of IOT devices being commonplace, the volume and velocity of the data comes into play. Most of this information is continuously streamed and may have to be analyzed in real or near real time. If the edge devices (the wearable devices in this case) in the contact tracing system were set up to undertake more of the processing workload, this could alleviate the amount of data pushed to the backend server for follow-on processing. This is similar to this contact tracing system. The contact tracing data will be streaming to data repositories daily, and the system has to be able to de-identify these streams of data if necessary. This operation also has to be able to scale to the enterprise level. This means that traditional static privacy rules may not be adequate, and "new adaptive privacy preservation and enforcement mechanisms are required" [42]. The de-identification method (i.e., differential privacy or k-anonymity) will have to change based on the stream and its content, and not rely on the ability to look at a static data set and remove the appropriate number of identifiers.

This concludes the description of the major privacy concerns that we identified surrounding the model of the DON's contact tracing system detailed in this thesis. The next chapter documents the application of the NIST Privacy Framework in a risk assessment for addressing the privacy concerns that we identified. The privacy concerns can be prioritized base on severity and likelihood of occurrence, then addressed in priority order in the development of the contact tracing system.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. RISK MODELING

A. PRIVACY RISK MANAGEMENT

The privacy concerns described in Chapter III and the application of the NIST Privacy Framework drive toward developing a privacy risk assessment of a system. Organizational leadership can then use this privacy risk assessment to make decisions about how to use information in an informed manner. As mentioned earlier, federal agencies are required to create Privacy Impact Assessments (PIA) for systems that collect, store, process, and distribute private data of employees and citizens. These PIAs effectively identify some problem areas that might arise when handling private data, but they are usually policy driven rather than technology driven. They do not directly link to implementing policy or technology controls to ensure privacy. Even when using the NIST Privacy Framework, PIAs are just one NIST SP 800-53 rev5 control of many when conducting a privacy assessment.

Risk assessment is another privacy control recommended by NIST SP 800-53 rev5. The fact that risk assessment is only one control can be misleading since risk assessment is such a large, underlying portion of system design, and it is an important aspect of determining the privacy afforded by a system. All privacy decisions affect the risk assessment. This point is more clearly articulated by the utility versus privacy tradeoff described in Chapter III. It is possible that instituting a small number of privacy controls can have a positive impact on our ability to analyze collected data; however, the risk of a privacy breach may increase—resulting in more risk to the organization and its customers and/or employees. This is why risk management is integral to assessing the DON Contact Tracing System and applying the NIST Privacy Framework.

Another area of focus that should be included in the risk assessment is threat modeling. Traditionally cyber threat modeling has been used by developers and designers to determine security risks to a system by identifying assets, attackers, and attack vectors. The NIST Privacy Framework does not explicitly identify threat modeling as a privacy concern even though there have been several organizations that have developed privacy

threat modeling frameworks, and two of those frameworks are discussed in the following sections.

The rest of this chapter will combine threat modeling with the NIST Privacy Framework in order to develop a rudimentary risk assessment model of the DON Contact Tracing System.

B. PRIVACY THREAT MODELING

Threat modeling is an invaluable activity in designing and evaluating software. It assists developers in identifying and understanding threats, vulnerabilities, and possible avenues of risk mitigation by trying to discover and model the actions of a malicious actor. The likelihood of attacks occurring is then combined with the potential damage they can cause, and this leads to a risk assessment of the system or app. Most threat modeling frameworks are intended to be used to reason about security risks. There also appears to be an assumption that good quality of security will translate into good quality of privacy. One of the oldest and most popular security threat modeling techniques is Microsoft's STRIDE (an acronym for Spoofing, Tampering, Repudiation, Information disclosure, Denial of service, and Elevation of privilege) which uses threat categories to discover security weaknesses that align with security properties within the CIA triad [46].

While threat model methods like STRIDE apply more to cybersecurity, that is not always sufficient for privacy in information systems. Privacy protection requires a more abstract way of thinking and a different threat model. That is not to say that privacy threat modeling cannot benefit from aspects of security threat models. For example, the STRIDE method begins with modeling the system by building a data flow diagram (DFD) [46]. This is also useful to privacy threat modeling; however, privacy entails more than just making sure that data is obscure to an external threat (i.e., confidential). Privacy comes with its own properties that need to be observed when threat modeling.

One of the recent developments in privacy-centric threat modeling are LINDDUN [47] and Cloud Privacy Threat Modeling (CPTM); an overview of both of these are given in this section. The DON Contact Tracing system utilizes a cloud system architecture; therefore, the threat model has to address this architecture in some way.

This concept of privacy threat modeling directly intersects with the NIST Privacy Framework. There are aspects of threat modeling baked into the NIST Privacy Framework, but the framework does not directly address threat modeling. That is why this chapter will create a privacy threat model for the DON Contact Tracing system. A well thought out threat model will facilitate the application of the NIST Privacy Framework.

The NIST Privacy Framework is being applied at the beginning of the System Development Life cycle (SDLC) for the contact tracing system, and that is also where threat modeling should initially take place regardless of the development framework used. If vulnerabilities can be discovered early in the development life cycle, developers can make architectural changes in a more timely and less costly fashion [46].

The rest of this chapter will discuss two different privacy-centric threat model methods, and then construct a basic threat model for the DON Contact Tracing system before applying controls from the NIST Privacy Framework.

1. LINDDUN

LINDDUN (a mnemonic for Linkability, Identifiability, Non-repudiation, Detectability, Disclosure of information, Unawareness, and Non-compliance) is a privacy-centric threat modeling methodology that uses a systematic approach in dealing with privacy threats. This approach involves first modeling the system using a DFD similar to STRIDE. Then threats are identified and mapped to the DFD using LINDDUN's threat taxonomy. Finally, threats are managed by choosing different mitigation strategies and privacy enhancing technologies (PET) [47].

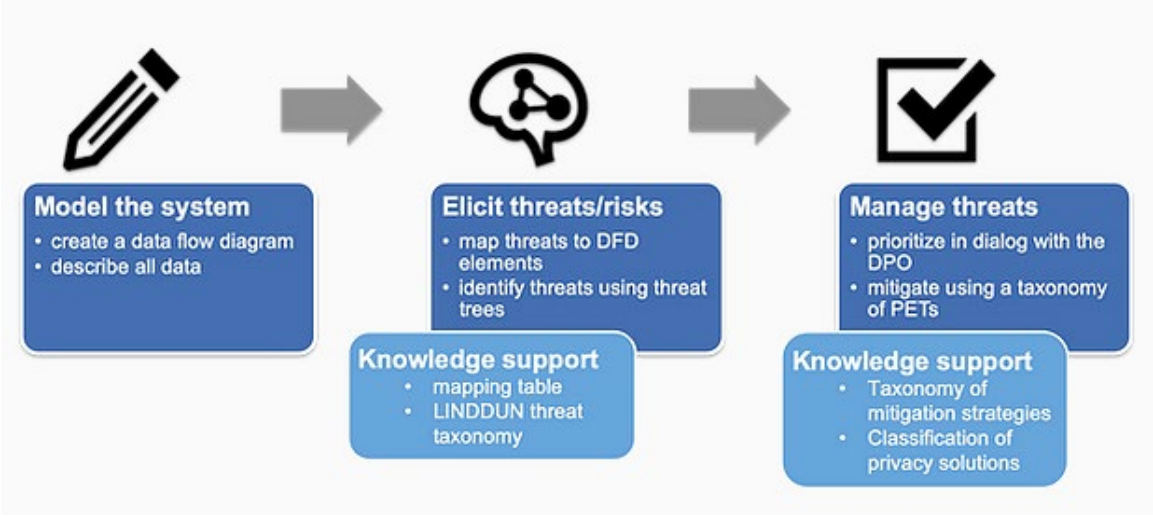

Figure 5. LINDDUN Methodology. Source: [47].

LINDDUN is not unlike the NIST Risk Management Framework (RMF) and the NIST Privacy Framework. The LINDDUN steps resemble the Categorize, Select, and Implement steps of the RMF. While the RMF has traditionally been more focused on information security threats, the NIST Privacy Framework isolates many privacy-centric aspects of designing or evaluating a system, much like the LINDDUN methodology.

Understanding the LINDDUN methodology will be helpful in creating a threat model for the DON Contract Tracing System because of the importance of the data flow aspect of the system and the importance of mapping threats to that dataflow. The data flow for the contact tracing system involves more than just local storage in a database at a command. The data in the system is being collected from individual users in the DON and then transmitted to and stored in the DON Enterprise Data Management System. The threat model needs to represent these and other data flows because there are risks associated with data in motion and at rest.

2. CPTM

There are data flows in the contact tracing system that involve cloud services; the threat model needs to represent cloud-specific privacy threats. The development of the Cloud Privacy Threat Modeling (CPTM) methodology was in response to the need for reasoning about privacy risks when data is handled by cloud services. CPTM was originally

developed to be used in accordance with the EU Data Protection Direction (DPD). DPD is tailored for use with EU privacy laws and regulations. However, the methodology can be applied in reasoning about privacy implications of applying cloud computing outside the EU [48].

Like LINDDUN, CPTM includes three main steps. These steps are as follows: identify the main entities to the cloud environment, describe the privacy requirements that must be implemented, and provide countermeasures for identified threats [48]. As shown in Figure 6, [48] also proposes extending the CPTM methodology steps to align with steps in the SDLC.

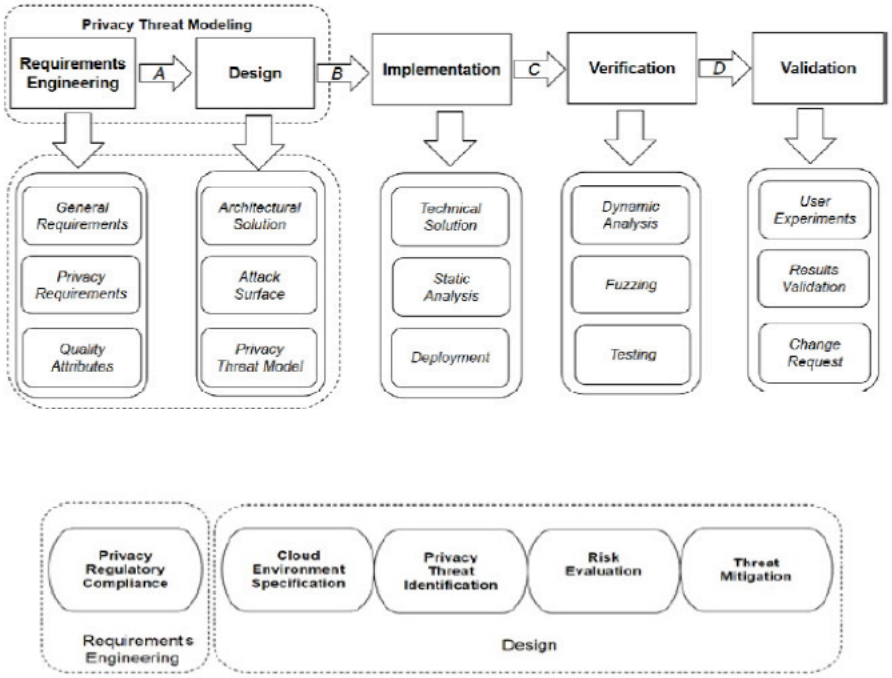

Figure 6. Cloud Privacy Threat Modeling (CPTM) and the Software Development Life Cycle (SDLC). Source [48].

These CPTM steps provide an overlap with some Core categories and functions in the NIST Privacy Framework, but with the CPTM methodology, the steps are more focused on a specific type of architecture and more aligned temporally when designing a system. The CPTM methodology reinforces the steps taken to build a privacy threat model for the DON Contact Tracing System.

C. CONTACT TRACING THREAT MODEL

It is important understand the distinction between security and privacy threats and where they overlap. While developing the privacy threat model, like with using LINDDUN and CPTM, threats to privacy are identified for the system at different stages in the data flow.

1. System Data Flow

Figure 7 shows at a high level of abstraction the entities and data flows for a hypothetical version of the DON Contact Tracing System. For a more detailed system diagram, see Chapter II.

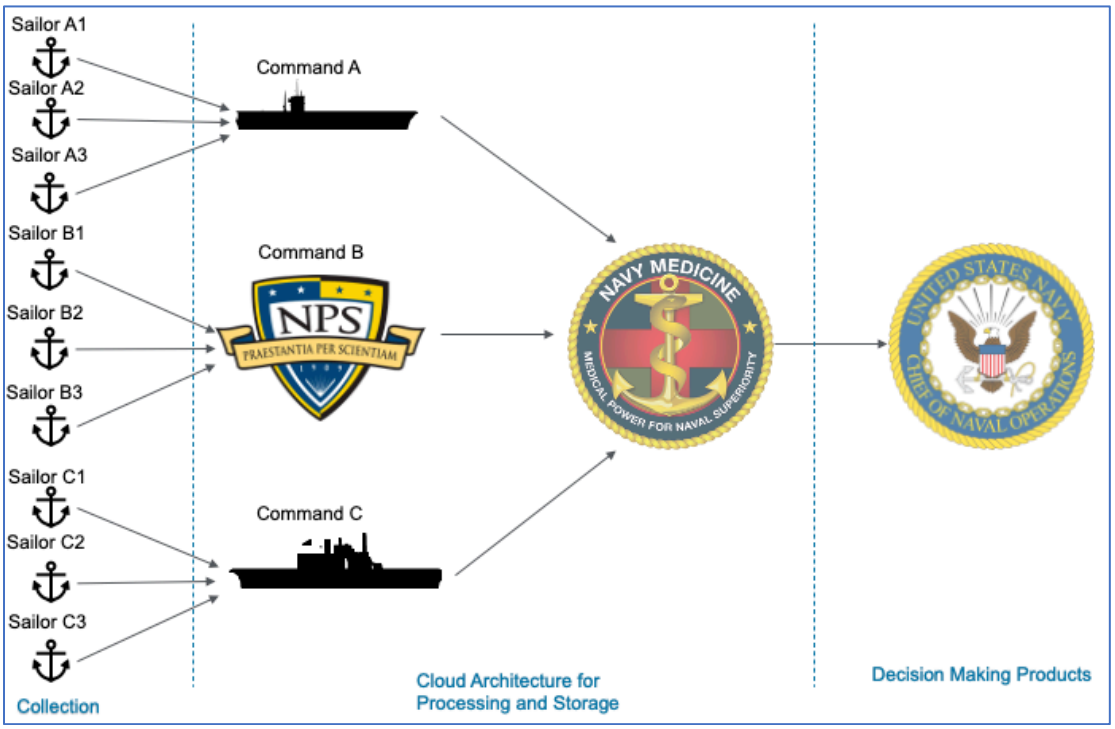

Figure 7. Contact Tracing System Data Flow

2. Threat Assessment

The following three sections show the privacy threats at different stages of the data flow.

a. Collection Stage

Table 1 lists ways in which private data is vulnerable during the collection stage. This is the stage when end user devices are transmitting Bluetooth identifiers to other end user devices and collection sites. It is important to note that physical security also comes into play in this stage.

Table 1. Threat Assessment during the Collection State

Threat	Vector	Mitigation	Impact
De-identification	Set up a collection device to capture Bluetooth identifiers and manually identify personnel to correlate to identifiers.	Physical Security Controls. Changing Bluetooth Identifiers.	Loss of trust in the system by DON personnel. Possible coercion attacks against de-identified individuals. Linkage attacks.
Social Graph Reconstruction	Use collected Bluetooth identifiers to infer connections between personnel.	Physical Security Controls. Encryption of transmitted data revealing proximity and exposure times.	Loss of trust in the system by DON personnel. Possible Organizational Chart Reconstruction. Embarrassment to individuals.

b. Cloud Architecture

Table 2 lists ways in which personal data can be vulnerable when being processed in the backend system.

Table 2. Threat Assessment during the Processing Stage

Threat	Vector	Mitigation	Impact
Social Graph Reconstruction	Analysts (authorized users) that have access to the data can discover social connections between personnel that should be private	De-identification Technology Access Controls Privacy policy involving data usage	Can manifest itself in the form of administrative repercussions for personnel (i.e., breaking curfew, being in unauthorized areas, fraternization) Embarrassment to individuals
Inadequate Policies	Data travels between databases under organizations that different privacy policies.	PIA Data sharing agreements Policy and Procedures	Legal ramifications if data is not properly IAW organizational regulations.
Unspecified Data Priority	Data that is of a lower priority may still be exposed to the same privacy concerns as higher priority data when it does not need to be	Policy and Procedures Data Tagging	Avoidable exposure of private data. Waste of time and man hours to sanitize and process lower priority data
Insider Threat (Intentional or Unintentional)	Unauthorized or authorized users accessing this data in the cloud that is not necessary for the execution of their duties	Policy and Procedures Access Controls De-identification Technology Auditing	Coercion and embarrassment of affected individuals Administrative repercussions against insider threats that could have been avoided

c. Decision-making Products and Announcements

Table 3 shows ways in which private data can be vulnerable based on a what kind of fleet readiness and decision-making products a commander requires. Reporting based on certain data private sets could also leak OPSEC information.

Table 3. Threat Assessment of the Dissemination Stage

Threat	Vector	Mitigation	Impact
Force Readiness	Decision makers announcing commands most at risk	Policy and Procedures. OPSEC	Mainly impacts the OPSEC of commands by revealing force readiness to adversaries Targeting of most at risk commands Possible false inject attacks to skew data
De-identification	Malicious actors can used products and reports to de-identify users by focusing on the most at-risk commands	De-identification technology	Coercion attacks Linkage attacks

D. CONTROL IMPLEMENTATION

This section focuses on the controls that should be implemented to mitigate the threats identified in the previous section. We are using controls from NIST SP 800-53 rev5 since NIST provides a mapping from the NIST Privacy Framework to controls in NIST SP 800-53 rev5. See Appendix B for a more detailed breakdown of the NIST Privacy Framework Core Functions and controls that can be applied specifically to this system.

The controls for this system can fit broadly into several categories which include Policy and Procedures (i.e., PIAs, Data Sharing Agreements, and Data Tagging), Disassociability (i.e., de-identification and linkage attacks), and Security Controls (i.e., Access controls, Encryption, Physical Security). The relationship between these categories is shown in Figure 8.

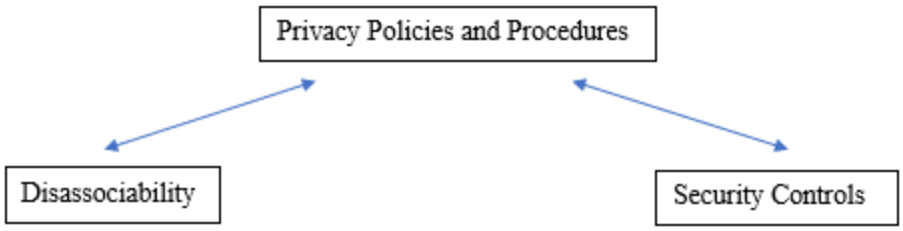

Figure 8. Control Categories and their Relationships to One Another

1. Policy and Procedures

NIST SP 800-53 includes a Policy and Procedures control as the first control in each control family of the publication. In fact, some of the Core Functions in the NIST Privacy Framework map to "all -1 controls," meaning that the subcategory being referenced is entirely mitigated by policies and procedures, regardless of the given control family from NIST SP 800-53 rev5 [6]. Additionally, other subcategories map to a Policy and Procedures control and that refers to a specific control family in NIST SP 800-5 rev5. For example, NIST Privacy Framework subcategory CT.PO-P1 maps to NIST SP 800-53 rev5 control PT-1 which addresses Policy and Procedures in regard to Personally Identifiable Information Processing and Transparency.

The PT-1 control states that "The risk management strategy is an important factor in establishing such policies and procedures" [6]. This is why policies and procedures should be a direct reflection of the risk assessment of a system. To that end, every organization and system will have different risk assessments which is why policy and procedures will be unique to those organizations and systems. This is why it is sufficient for NIST 800–53 rev5 to not be highly detailed in its guidance on implementation of policy and procedures. It is important for the framework to emphasize this step because this step lays the foundation for legal, regulatory, and technical controls needed for privacy assurance.

2. Disassociability

Aspects of disassociability could be included under Policy and Procedures since organizations should have a standardized procedure for de-identifying data, and that should be based on what the organization considers an acceptable risk of re-identification and

linkage attacks. Disassociability to include de-identification also has a technical component which is why it is included in a category separate from Policy and Procedures.

The NIST Privacy Framework provides a Category called Disassociated Processing (CT.DP-P) which provides data processing solutions that "increase disassociability consistent with the organization's risk strategy to protect individuals' privacy and enable implementation of privacy principles" [1]. This encompasses more than just de-identification and linkage attacks, but ultimately this represents technical privacy-by-design choices and controls that need to be applied in order to provide privacy assurance.

Many of the controls in this Disassociated Processing Category relate to design and architecture (see Table 4). This circles back to one of the first considerations of the system – whether or not a centralized or decentralized system will be used. This design choice, however, is not specifically mentioned in the NIST SP 800-53 rev5 controls. These kinds of decisions are left up to the developers to make during system design. The NIST standard provides guidance instead of being prescriptive.

Table 4. Controls in the Disassociated Processing Category Related to Architecture and Design. Adapted from [1], [6].

NIST SP 800-53 rev5 Control ID	Control Name
PL-8	Security and Privacy Architecture
PM-7	Enterprise Architecture
SA-8	Security and Privacy Engineering Principles
SA-17	Developer Security and Privacy Architecture and Design
SC-42	Sensor Capability and Data
CM-6	Configuration Settings

The NIST Privacy Framework treats de-identification as a product of how the system is designed and implemented; it also addresses this issue directly with control SI-19 (De-Identification). De-identification is singled out because if the de-identification policy and corresponding technology do not sufficiently protect user privacy ("sufficient" being a relative term here referring to what an organization and its customers deem acceptable), the privacy risk can grow outside what is considered by the organization to be tolerable. This will be discussed more in the next chapter.

3. Security Controls

While not the focus of this research, security controls play a role in privacy assurance. The extent to which security controls should be included in the privacy conversation is discussed in the next chapter. The security controls of a system are usually technical in nature, but when it comes to privacy assurance, these security controls need to be integrated into the privacy policies and procedures of an organization that manages data to be kept private. This section highlights that the threat model identifies security-based controls that need to be implemented and included in an organization's privacy policies and procedures.

The Protect Function (PR-P) in the NIST Privacy Framework is almost entirely focused on security-based (CIA) controls that provide a level of privacy assurance. This includes controls that enhance authentication management, access control, and transmission confidentiality; and all of these types of controls are required by the contact tracing system to some extent (see Table 5). The NIST Privacy Framework in conjunction with NIST SP 800-53 rev5 addresses these security aspects as overlapping with privacy, but they leave it up to the designer of a system to be able to integrate the security and privacy controls in a way that ensures both security and privacy.

Table 5. Selected Security Controls that Enhance Privacy. Adapted from [6].

NIST SP 800-53 rev5 Control ID	Control Name
PL-2	System Security and Privacy Plans
IA-2	Identification and Authentication (Organizational Users)
IA-3	Device Identification and Authentication
IA-8	Identification and Authentication (Non-Organizational Users)
AC-24	Access Control Decision
SC-8	Transmission Confidentiality and Integrity

E. CONTACT TRACING RISK ASSESSMENT

The use of privacy threat modeling and privacy control selection are parts of a risk management strategy. NIST provides guidance for risk management in the form of its Risk Management Framework laid out in NIST SP 800-37 rev2 [19]. This publication provides guidance for security and privacy risk management, although only the NIST Cybersecurity Framework has been integrated into the RMF cycle. This may be due to how new the NIST Privacy Framework is at the time of this writing. Regardless, NIST has established guidance on how to perform security and privacy risk management in NIST 800–37 rev2, and the NIST Privacy Framework also includes mappings to risk assessment controls in NIST SP 800-53 rev5. In fact, NIST SP 800-53 rev5 contains an entire family of controls called Risk Assessment [6].

NIST also provides guidance specifically on privacy risk management in its Internal Reprot 8062 [3]. This report goes into detail about engineering and privacy risk modeling, and these area contribute to NIST publishing the NIST Privacy Framework in 2020.

Risk Management Strategy is also a Category of the NIST Privacy Framework (GV.RM-P), and it provides mappings to controls that help with evaluating risk tolerance by way of Risk Framing (PM-28). Every organization will have its own unique risk tolerance that will lead to its own risk management strategy, but there are a few principles of risk management that all systems have in common. Since the exact details of the DON Contact Tracing system have not been established, it is only possible to make a general risk assessment of the system in its assumed configuration.

There are two main risk management questions to answer for this system (as well as any system with privacy concerns). The first is whether or not the assessed risk is greater than the tolerable risk. Secondly, is the NIST Privacy Framework and its other associated references able to lead an organization to the answer to the first question?

1. Determining Risk Based on Threats

To answer the first question above requires a calculation with several variables shown below in Figure 9. This privacy risk equation is taken from the NIST Internal Report 8062 and shows the interaction of concepts realated to privacy risk [3].

Identify Threats and Vulnerabilities x ***Determine Likelihood of Occurrence*** x ***Determine Impact of Occurrence*** $=$ ***Risk Assessment***

Figure 9. Risk Assessment Equation. Adapted from [3].

This equation uses variables taken from the Risk Assessment control family of NIST SP 800-53 rev5 and NIST Internal Report 8062. This calculation can be more fine-tuned, empirical, and possibly automated in order to facilitate decision-making for the Authority to Operate (ATO) official, but the fundamental principles still stand.

A rudimentary example of using a threat we identified during threat modeling is shown in Figure 10.

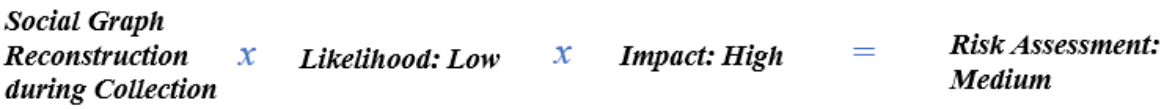

Figure 10. Risk Assessment Equation Example

For this threat to be a privacy vulnerability, an attacker would have to intercept the user's Bluetooth beacons as well as the time and distance calculations determined by the wearable devices. This would require an adversary to obtain physical access to the command (within range of the Bluetooth beacons) in order to set up a collection device to intercept the beacons being exchanged and the data being uploaded to the collection hubs. The adversary would also need to have the ability to decrypt the transmissions since the transmissions containing time and distance calculations will likely be encrypted when transmitted to the collection hubs. Given this information, this makes the likelihood of this attack rather low.

The impact, however, could be determined to be high because if an adversary gets access to a user's social interactions via this contact tracing system, there could be several harmful outcomes. The adversary could use this information as part of a coercion attack if they can determine that the user is in the vicinity of an area or a person that they are not supposed to be near (e.g., restricted areas and/or fraternization). The adversary could also glean critical OPSEC information about a user if they can use the beacon information to determine that the user is always in close proximity to other identified members of the command. For example, they could determine that a user is always in close proximity to the CO or XO, and that could make this person a valuable target for surveillance, coercion, or even physical attack.

Each organization will be able to come up with its own metric on how to make the final risk assessment, but in this example, the risk assessment would probably be determined to be Medium. That is based on the likelihood of occurrence being low, and the impact of occurrence being high.

The next two components to the risk assessment involve the organization's risk tolerance and the risk mitigation controls. The risk tolerance will differ between organizations so this paper cannot make an accurate determination of what the DON's risk

tolerance will be. The risk tolerance may depend on external factors like how damaging the COVID-19 outbreak is or the sensitivity level of personnel at a command. The command may be willing to assume additional risk if the COVID-19 is significantly affecting vital operations. Another reason to assume more risk would be if the command in question conducts fewer sensitive operations. The social interactions of personnel at a command like NPS or USNA may not be as sensitive as the interpersonal interactions between sailors to a Carrier Strike Group.

Regardless of the reasoning, system designers need to compare the organization's risk tolerance to the risk assessment calculated above. If the risk assessment is greater than the risk tolerance, designers need to take action. They can either improve or modify the controls, or in extreme cases, they can determine that the system implementation is not worth the risk. The following snippet of pseudocode in Figure 11 emulates how a system designer should think about risk assessment versus risk tolerance.

```
while (riskAssessment(threat, likelihood, impact) > riskTolerance)
{
      modify controls;
      update likelihood;
      update impact;
}
```

Figure 11. Risk Assessment versus Risk Tolerance

This section serves to illustrate the general method for risk modeling needed for this contact tracing system. It is important because the risk modeling flows from the threat modeling, and this eventually leads to the application and modification of privacy controls, technical or otherwise. Those controls can then be integrated into privacy policies and procedures, which is paramount in privacy assurance.

2. How Does the NIST Privacy Framework Operationalize a Risk Assessment?

The ultimate goal of an organization using the NIST Privacy Framework should be the ability to develop a risk assessment that contains a list of risks and possible controls

that can be presented to organizational leadership and consumers alike. With the information derived from the framework, the entity that is approving the use of the system should be able to make an informed risk decision about whether the system should be operational in its current configuration. With the DON Contact Tracing system there is an inherent risk to mission accomplishment, whereas in the commercial sector, the risk is more monetary in nature. The question then becomes whether or not the NIST Privacy Framework can be operationalized in a manner that properly identifies the risks in the contact tracing system and provides acceptable controls to manage those risks.

As shown in the previous sections, the threats and vulnerabilities that were identified in the contact tracing system could be matched to controls in NIST SP 800-53 rev5 that would be expected to mitigate those threats and vulnerabilities to some degree. The process of identifying threats and vulnerabilities is far more nuanced and can be found mainly nestled in the Identify Core Function of the privacy framework. If a developer of the contact tracing system were to go down the list of Categories and Subcategories of the Identify Function (similar to what was done in Appendix B), he or she would come across some broad sweeping privacy aspects of the system in question that need to be taken into consideration. This would be aspects like identifying data actions, data elements, data ecosystems, and the processing and purpose associated with these aspects. These Subcategories are not sufficient to lead a developer of the contact tracing system directly to the specific threats. For example, the Disassociated Processing Category (CT.DP-P) is located under the Control Function when it should also be located under the Identify Function. Disassociability is one of the major concerns with this contact tracing system, and a huge privacy concern in all systems that collect and process private data. The organization's policy on Disassociated Processing should be clearly addressed in the Identify Function before considering controls for it in the Control Function. The disassociability concerns should also be addressed in the same Functions and Categories that also address Risk Assessment.

Risk Assessment is another Category under the Identify Function that, while necessary and important to developing a system like the contact tracing system, seems hidden within the framework. The entire framework should be considered as a risk framework for privacy. If an organization does apply the NIST Privacy Framework to a

system, and that organization is not any closer to an effective risk management strategy, then the framework has failed. Specifically, the Risk Assessment Category under the Identify Function probably deserves a bigger role in the framework. The Risk Assessment Category mentions the concepts that we described in this chapter in terms of creating a risk assessment, but it does not provide a tangible method to calculate this risk assessment. Most of that detail is left to the individual entities designing or assessing an information system, and that is completely acceptable due to the unique nature of most information systems. The main point is that risk assessment and risk management should be the goals of the entire framework and not just a part of it, and that should be reflected in how the framework is presented.

Operationally, the NIST Privacy Framework is not as linear as it appears. It appears to take an approach that involves identifying concerns, creating a risk assessment, and implementing controls (mainly security controls). While the cyclic nature of these processes in likely implied by the framework, it would be valuable to emphasize the cyclic structure of risk management. NIST should constantly be referencing its famous RMF guidance while talking about its privacy framework since they are both ongoing cyclic processes that require users to revisit aspects of the assessment for reevaluation. Figure 12 shows the cyclic nature of risk assessment by mapping categories in the privacy framework to steps in the RMF.

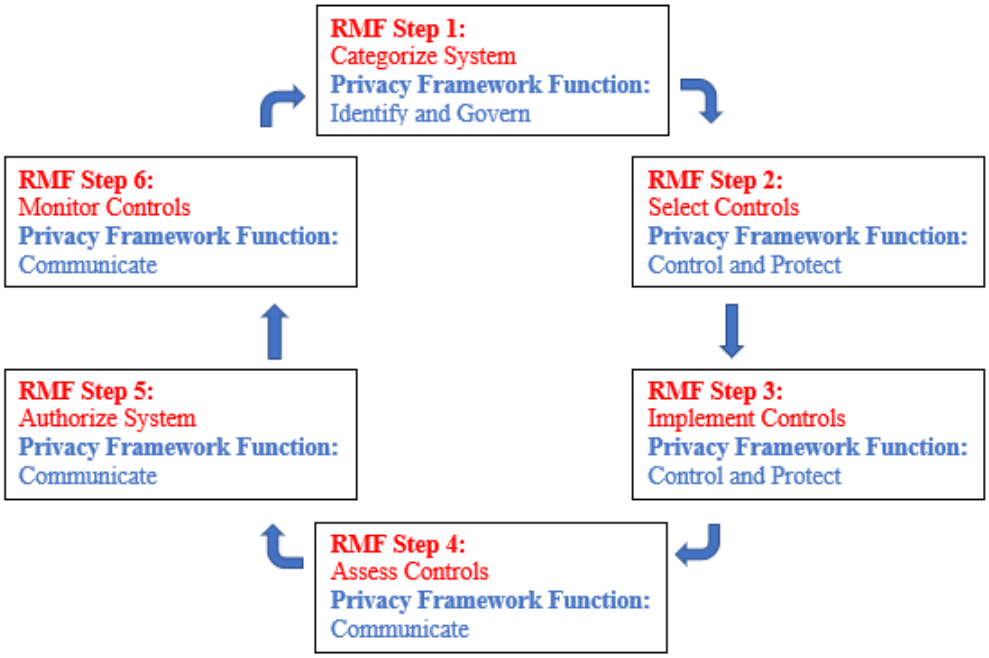

Figure 12. NIST RMF and NIST Privacy Framework Mapping

NIST states in its "Roadmap for Advancing the NIST Privacy Framework: A Tool for Improving Privacy through Enterprise Risk Management" that the "privacy domain lacks development and uptake of uniform concepts of privacy risk assessment, including specific risk factors, as well as more in-depth guidance and tools for assessing privacy risks" [22]. This acknowledgement goes to show that NIST understands that privacy risk assessment is a challenging endeavor right now, and many organizations do not know how to integrate privacy risk management into their current risk management approaches. In the context of the DON Contact Tracing system, the privacy framework can be operationalized to a degree, but it is up to skilled technologists and policy makers to understand how the priorities of their individual system before just going down the privacy framework like it is some type of compliance checklist. NIST has provided an invaluable resource to use in privacy risk assessment, it just needs to be placed in the right context when designing or assessing an information system.

Ultimately, the NIST Privacy Framework does accomplish its purpose which is to help organizations manage privacy risks. The degree of effectiveness of that claim is yet to be seen as the NIST Privacy Framework is still relatively new (it was released in January

2020). The NIST Privacy Framework gives developers and assessors a list of problems to consider and provides a rudimentary roadmap to controls that may solve those problems. Although the NIST Privacy Framework is not mandatory for federal agencies, ideally one would imagine that a proper application of the framework should be enough to certify that a system is within the limits of acceptable privacy risk for an organization. This also might not be the case since there are nuanced policy and technology aspects that may need to be considered when applying this framework. The next chapter goes into more depth about some of the pressing issues regarding privacy risk in the contact tracing system that were revealed by the risk modeling in this chapter, including disassociability, privacy policy, and security controls that require a deeper understanding about how they are applied within the system.

V. DON CONTACT TRACING SYSTEM ASSESSMENT AND RECOMMENDATIONS

This chapter uses the findings from the previous chapter on privacy risk modeling to provide an assessment and make recommendations about the DON Contact Tracing System. These recommendations are meant to be broad and somewhat overarching since the system is not currently in place and the exact details of the system are not currently known; however, these recommendations must still be considered when implementing an information system that collects, processes, and transmits private data in a manner such as the one described in Chapter II.

A. THE IMPORTANCE OF SOUND PRIVACY POLICIES AND PROCEDURES

Above everything else, the threat and risk modeling of the contact tracing system brought to light the need for a well-developed privacy policy that takes into account privacy risk when implementing appropriate technical controls. Digital privacy revolves around the responsible use of data, and having tailored privacy policies and procedures in place ensures that organizations are taking the most effective steps toward being responsible with data about individuals. Privacy policy can be used to inform users of a system or app about how data about them will be collected, stored, and processed. Privacy policy can also specify how users can interact or control these data actions when they involve their personal data. Most privacy policies are designed in some way that complies with a legal regulation such as the GDPR or HIPAA. The DOD, DHA, and DON all have privacy offices that provide their own broad privacy policies and procedures that are applicable to their respective organizations, but since every system is unique, there should be more guidance on the data actions of the DON Contact Tracing System.

As mentioned in Chapter III, all DOD systems that collect, store, or process PII require a PIA. PIAs are a useful tool mainly from a legal standpoint, but information systems also require detailed procedures for ensuring technical privacy controls are implemented in such a way that risk is mitigated to the maximum extent possible. If a privacy policy does not properly take into account the appropriate risk management factors

(discussed in Chapter IV) when implementing controls, even the most advanced and well-tested controls may be ineffective at protecting privacy. For example, having a policy that states that all contact tracing data will be de-identified before being uploaded to the DON enterprise cloud architecture may be an ineffective policy. As it is phrased, this policy would not take into account the loss of utility if the data were uploaded with missing identifiers. This policy also does not account for the risk in re-identifying the data. The ability to re-identify data is always present (see Section B), and the privacy policy and procedures should be clear about what specifically should be done in order to ensure data is de-identified and to what extent does it need to be de-identified (i.e., which identifiers or quasi-identifiers need to be removed in order to fall below the risk threshold established by the organization?).

Another consideration for framing DON privacy policy for its contact tracing system is the difference between the DON mission focus and the focus of commercial concerns. The DON is less concerned than businesses about the monetization its data in terms of selling data to third parties; the DON does not have the problem faced by businesses of losing customers due to privacy policy or breaches of that policy. Instead, the DON can center its attention on the individual rights and dignity of the users and more broadly on which privacy policy will best facilitate accomplishment of the DON's mission. The privacy policy should instill trust in its users assuming that the second order effect of this trust will lead to a mission-ready workforce, for which an important precondition is keeping the spread of COVID-19 at bay. The following sections delve deeper into an areas of concern identified in Chapter IV.

B. A POLICY FOR DISASSOCIABILITY

Disassociability, a term taken from the NIST Privacy Framework, is used in this paper to include both de-identification and anonymization. Both of these terms sound similar, but sometimes they are described as having slightly different meanings. According to [49], de-identification specifically involves the explicit process of removing identifiers from a dataset, whereas anonymization is not a method or process but more of an end goal of a combination of methods or processes. In other words, anonymization is the desired end result of de-identification if done properly. NIST claims that these terms are sometimes

used interchangeably, but NIST also acknowledges that sometime anonymization denotes an unreversible type of de-identification [50].

This research mainly focuses on de-identification of data over anonymization for use in applying the privacy framework to the assess the risk posed by the contact tracing system. The term de-identification is less ambiguous than anonymization. De-identification is also integral to performing data actions on PII.

NIST has conducted extensive research on de-identification of personal data in governmental datasets. The results of that research needs to be used to inform the privacy-related requirements for the DON contact tracing system [43], [50]. The focal point of the NIST-conducted research was on making these datasets releasable for research while simultaneously protecting the privacy of the citizens whose data is in these datasets. With the contact tracing data, if it is re-identified to be stored with other medical data about the users, this data will need to be de-identified again in order to make it available to research entities that are either internal or external to the DOD.

The research conducted on de-identification by NIST is a valuable reference for all things surrounding de-identifying data at the federal level, but it is also important to take into consideration the stance on de-identification by two of the most prominent pieces of privacy, the GDPR and HIPAA. Both the GDPR and HIPAA have different standards of de-identified data. The GDPR takes a spectrum approach and defines levels of de-identification as Identified, Identifiable, Article 11 De-identified, and Anonymous/Aggregated [51]. This is a useful approach in theory because it has been shown that re-identification is possible to achieve by leveraging MI/AL techniques [52]. While this does have implications for classifying de-identification standards within the different GDPR de-identification levels, it does show that the GDPR acknowledges that de-identification is not strictly binary like one would see with HIPAA de-identification standards [53]. When applying the GDPR spectrum-based approach to de-identification, it gives more justification to the risk-management approach of privacy which in turn implies that the levels of de-identification need to be continually updated based on advances in computing that change the risk calculation.

HIPAA takes a more binary approach and considers de-identified data to no longer be private data, and therefore no longer falls under HIPAA protection. This means that data that has been properly de-identified in accordance with HIPAA standards can be shared with research organizations that are not covered entities. To the credit of HIPAA, the legislation does provide a robust standard for de-identification. This includes the following two methods: the Safe Harbor method, where 18 types of identifiers are removed and the Expert Determination method where an expert assesses whether information has been de-identified below a particular risk threshold based on the application of statistical and scientific principles [49].

Understanding the HIPAA standard of de-identification is important for the contact tracing system, not only because the contact tracing data might be linked to HIPAA data at some point in the data life cycle, but also because of how the HIPAA means of de-identification will affect the utility of the data. Contact tracing data does not fit squarely into any of the 18 types of identifiers that are removed in the Safe Harbor method [53]. It has already been proven that valuable social information can be determined by Bluetooth data, which in this case is the proximity data provided by the contact tracing system [38]. This makes it vital that the DHA and DON determine a standard by which contact tracing data can be shared within the DOD enterprise while limiting the risk of re-identification and while still maintaining an acceptable level of data utility.

The discussion about GDPR and HIPAA standards on de-identification serves to illustrate the point that disassociation needs to be dealt with in much more complex terms than the NIST Privacy Framework leads a user to believe. It involves more than just stripping names from spreadsheets and encrypting data. Privacy assurance, as stated in Chapter IV, is how policy and procedures handle the intersection of disassociability and security controls; therefore, these policies and procedures need to be tailored to a specific system like the contact tracing system. The following paragraphs present some recommendations specific to the DON Contact Tracing System that should be weighed when developing a policy or standard for disassociation of user data.

First, the DON needs to acknowledge that the contact tracing data is sensitive information in terms of disassociability regardless of whether or not it is directly associated

with identifiers or quasi-identifiers. Therefore, it should be protected to the greatest extent possible when stored or transferred within the system. Even though identifiers are not linked to the contact tracing data when collected, Bluetooth proximity data can be used to infer social graphs [38]. There is the aspect of deniability since it is difficult to associate a Bluetooth Identifier to a user with 100% certainty, but the social graphs that can be inferred still represent a threat to privacy and to command OPSEC since operations can also be inferred.

The DON needs a plan for dealing with mappings to Bluetooth Identifiers in the system. The original configuration has several nodes where the contact tracing data can be re-identified or linked to a user directly. The first time would be when the individual commands issue the devices to employees. That command will maintain a mapping in order to keep track of the devices, and this will be stored in the cloud services provided by the device vendor. The other location where users will likely be re-identified is in their medical records at BUMED. This has implications for data use since the contact tracing data is not considered PHI before it reaches BUMED. It also raises concerns about who is able to access the contact tracing data at each command. This leaves room for abuse if a Division Officer or Division Chief Petty Officer can access the collected data and make a determination about whether or not a Sailor has been spending enough time in his or her assigned workspace. To remedy these issues, each command should have Navy Medical issue the contact tracing devices and have all of the contact tracing data route directly to BUMED databases (see Figure 13 for a visual representation of the proposed chain of custody of contact tracing data). This will remove the temptation for misuse as well as simplify the data flow. With all the information going directly to BUMED, the data will begin as PHI before any entity has had a chance to analyze the data. This also requires that the contact tracing device vendor provide a system that does little to no minimal backend processing before pushing the data to BUMED.

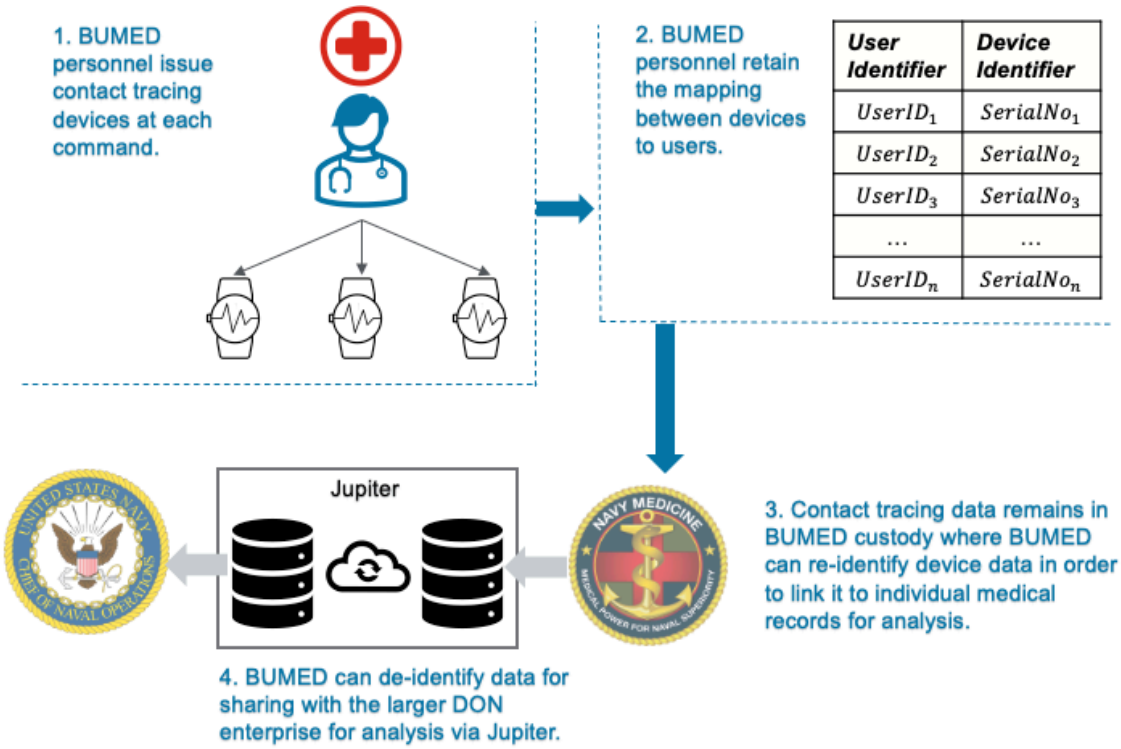

Figure 13. Chain of Custody of Contact Tracing Data through BUMED

Pushing all the contact data directly to BUMED will make the data PHI and subject to HIPAA regulations regarding de-identification. This is where the most crucial de-identification decisions need to be made as this contact tracing data has the potential to be shared within the DON Jupiter system for analysis and reporting purposes. There are two options in this scenario, and they are either sharing the data after it has been de-identified appropriately or conducting all analysis under the purview of BUMED and only sharing the results of the analysis in a manner where there is a low risk of reverse engineering the result to re-identify either organizations or individuals.

In the first case, sharing the contact tracing data will probably require the Expert Determination method since this method requires a more blatant risk assessment method in appraising the risk of re-identification. Contact tracing data is linked to privacy more so than characteristics like medical conditions because relations, actions, and locations can be inferred from the contact tracing data. This means that contact tracing data is not only a privacy risk to individuals, but also to commands and the military at large. For this reason,

an expert(s) should be required to determine the method of de-identification required before sharing the data by conducting a tailored risk assessment for this specific type of data (for more on making risk assessments, see Chapter IV).

In the second case where data is processed and analyzed within BUMED, researchers need to take caution when sending analyzed data downstream in the data path because certain analyzed data may have the potential to leak identifying information about individuals or commands. Risk of follow on uses of data is not considered in the NIST Privacy Framework, but it should be applicable in this contact tracing system. For example, a detailed analysis of a command's COVID-19 risk mitigation strategies using contact tracing data could reveal private information about the command or individuals at the command. If a ship that has ten Sailors that work in the same department, seven of them contract COVID-19, and those seven Sailors were not social distancing effectively enough according to their contact tracing data, it would give anyone a 70% chance of guessing who has COVID-19 in that department as long as they can recover a list of the 10 people working in that department. Even though the personal information has been removed from that report, the results of the analysis may lead to re-identification. This is a primarily a threat to OPSEC since adversaries may be able to leverage this information they have on Sailors and commands. Maybe the risk of potentially exposing those seven Sailors' diagnosis is a risk the DON is willing to take in order to curb the spread of the disease, but this is a determination that needs to be made by experts that understand risk assessment in the privacy space as it concerns force readiness and OPSEC.

These recommendations all come back to policy and procedures. The technology for de-identification is not perfect and every de-identified data set is at some risk of re-identification. This is why it is incumbent on the part of privacy experts to be involved in the process of risk assessment and policy making when it comes to disassociation of these contact tracing data sets. In other words, there are privacy requirements for disassociation. Privacy experts must also represent the interests of privacy as system engineers formulate tradeoffs among the dependability attributes of the system, such as safety and security.

C. SECURITY REQUIRMENTS FOR PRIVACY

It is well known in privacy research circles that security can exist without privacy, but privacy may not be able to exist without security. The risk assessment documented in Chapter IV reinforces that need for security in order to maintain privacy of an employee's contact tracing data. Security controls are necessary in the contact tracing system in order to ensure privacy especially in the areas of data collection and access control when the data is stored.

In the area of access control, it was already recommended earlier that the contact tracing devices and their user device mappings only be overseen by the medical staff at each command. While this aids in disassociability, it is also a recommendation that limits access control and limits the number of personnel that are allowed to view this contact tracing data and are able to re-identify it. Minimizing the number of personnel with access to the re-identified data aligns with the principle of least privilege which in this case is seen as both a security and a privacy control.

The other recommendation that comes from the risk assessment of this contact tracing system is the need for confidentiality commensurate with the level of sensitivity of the contact tracing data. The collected data is sensitive regardless of whether or not it has been re-identified. This is why cybersecurity measures, such as employing end-to-end encryption, should play a significant role from the point where the data is collected to the point where the data is secure in BUMED and Jupiter databases. The cybersecurity measures need to ensure it is difficult for an ease dropper or man-in-the-middle to identify patterns in the transmitted data that might reveal serial numbers of the devices or other sensitive information like proximity or time data.

Thankfully, this endeavor is already being undertaken by engineers at Naval Sea Systems Command (NSWC) Crane Division. NSWC Crane was tasked with identifying security flaws in any contact tracing system that the DON is researching for procurement. This is essential, not only for security, but also for privacy; however, ensuring secure communications in a system in not enough for putting the privacy stamp of approval on the system. This is evident by the previous section on disassociability recommendations.

Security controls cannot account for all the privacy assurance in an information system, and in many cases, security is used as a substitute for privacy.

A recommendation for DON engineers is to conduct dedicated privacy testing (in addition to security testing) in regard to information system procurement. This would push the DON to place a higher weight on the technological controls for privacy while combining them with the regulatory/policy controls that the DON already implements with its privacy program. Privacy compliance is moving closer to the front of conversations about personal data much in the same way cybersecurity has taken a prominent role in the conversation surrounding national security. It would be beneficial for the DON to get ahead of this movement and begin establishing a division in the privacy office that deals with privacy technologies and assessments of privacy technologies.

Additionally, to aid the privacy assessment of this contact tracing system, it will require a test bed to ensure privacy is being preserved during operations. To its credit, the DON was planning on using the United States Naval Academy as their testing grounds for the system. It was to be conducted on a voluntary basis, and its purpose was to test the operation of the system. While this type of real-world testing is valuable to ensure proper operation of the system, that also includes ensuring the system is secure and privacy preserving. Based on the leadership's description of the contact tracing system being tested, many of the midshipmen were apprehensive to take part in the test since their biggest concern was privacy. The midshipmen's concern centered on the ability of the system to track their movements in such a way that inferences about their activities could lead to administrative repercussions. Another concern is that the privacy controls in the pilot system might be ineffective. Security has little to do with alleviating these fears because access controls and confidentiality controls that are intact can still lead to breaches of the individual's privacy. This is another instance where privacy policy takes the lead, and the mixture of a risk-informed privacy policy with appropriate and tested disassociability and security controls should be able to alleviate these concerns and privacy gaps under a chosen risk threshold. Removing fear is in part related to addressing uncertainty and being able to inform users about actual versus perceived risk.

The reaction of these apprehensive midshipmen is the reaction that the DON should want to see. The workforce itself needs to become more privacy conscious and provide feedback like this to system developers. This is why pilot studies, testbeds, and stakeholder-feedback sessions are vital to achieving privacy goals and objectives of the DON. They provide the foundation for making a risk assessment like what was documented in Chapter IV. The system being implemented can be tailored to fit the security and privacy needs of the DON based on that risk assessment.

D. THE IMPORTANCE OF RISK ASSESSMENT

In addition to providing justification for risk-informed policies and procedures surrounding disassociability and security, the risk modeling in Chapter IV revealed the need for the DON, first and foremost, to quantify its risk tolerance when assessing or designing an information system like this contact tracing system. Many manufacturers of digital contact tracing systems will make claims like "Privacy Guaranteed" or "Completely Secure" which any technologist should see that as a marketing ploy aimed at convincing customers that this company has taken sufficient steps to ensure security and privacy. A knowledgeable designer or engineer should recognize that there is no such thing as perfect security or perfect privacy in practice. Companies will claim that removing identifiers ensures one hundred percent privacy or that providing encryption will also guarantee privacy. This research has already mentioned that de-identification is not a perfect privacy-preserving solution, and that security controls do not address all aspects of privacy.

This is where risk modeling and risk assessment comes into play. Reverse engineering a contact tracing system in order to find security flaws is vital to the process of procurement and ensuring privacy as well, but it is not enough in ensuring overall privacy in the system. This comes from first establishing the threats/vulnerabilities, their likelihood, and their impact while simultaneously determining the risk tolerance of the organization. Without this assessment the designer of a system would not be able to determine the extent to which a chosen control will lower the risk, and they will also not have a risk tolerance to compare it to. In turn, there would be no informed basis for choosing which controls from the NIST Privacy Framework to implement.

This brings the conversation back to the companies that sell cyber snake oil. Policy makers and system developers need to ignore those claims and see privacy as more than just a technical solution or a policy solution and more as the intersection of technical mechanisms and policy implementations that are backed by risk determinations. Whether or not the NIST Privacy Framework is the best tool to make these determinations is discussed in the next chapter.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. NIST PRIVACY FRAMEWORK ASSESSMENT AND RECOMMENDATIONS

The NIST Privacy Framework is certainly a step in the right direction in the realms digital privacy and software and system engineering. By NIST's own admission, the real value of the NIST Privacy Framework is, or at least should be, facilitating communication between technologists and c-suite and opening a dialogue about privacy and its impacts when designing systems, products, and services [1]. This dialogue brought on by this framework should inevitably lead to a determination of risk tolerance and a risk assessment for the organization in regard to privacy.

On this path to a privacy risk assessment, there are several aspects of the NIST Privacy Framework that need to be addressed and either emphasized or clarified in order to better illuminate this path. This chapter revisits the privacy and security relationship as it pertains to the actual framework and whether or not the cybersecurity and privacy frameworks should be separate documents. It also discusses the checklist nature of the framework and how effectively this translates to a privacy risk assessment. Finally, this chapter dives into the privacy framework as an identification tool and explores some of the constraints regarding its application.

A. PRIVACY AND SECURITY FRAMEWORK INTEGRATION

The last chapter detailed the need for security controls in the contact tracing system in order to ensure privacy. Privacy can and does require security centric controls, but just implementing security controls does not always ensure a tolerable level of privacy risk. Security only ensures that unauthorized users cannot view data that they are not supposed to be able to access. It does nothing to protect against inferring information from data that authorized users have access too, and security sometimes does not control what users can do with data sets after they are granted access to them (i.e., being able to re-identify individuals in an anonymized data set that has been released to the public).

The NIST Privacy Framework upholds this relationship between Privacy and Security, by including a Protect Function that focuses on data security, but this does not quite capture the full essence of the privacy-security relationship. In all likelihood, NIST

might have included this Core Function to ensure that users of the framework know that security is necessary for privacy in many instances. The issue here is that there already exists a NIST Cybersecurity Framework that includes this same function as well as similar subcategories [20].

The two frameworks intentionally have a similar structure so that they can be used concurrently to analyze risk for systems that have requirements for privacy and security, and this raises the following question: Should there be two frameworks or just one? The answer to this question is based on the degree to which a practitioner believes privacy and security should either be isolated or coupled, and this is the crucial dilemma that needs to be reevaluated by NIST.

The prevailing thought of this research is that the privacy and cybersecurity frameworks should be included under one framework that developers and engineers can tailor to the system depending on the level of privacy or security needed for that system. In this setup, it is harder to completely decouple them, and it is easier to leverage the concepts together toward a common goal. This is not to say that security needs to be considered with every application of privacy. This was seen in the previous chapter when referring to disassociability and how security sometimes can have little impact on this aspect of privacy. The combining of the two frameworks would represent how the two concepts are related but the different Functions and Categories would be able to represent how security and privacy challenges can be isolated in certain contexts. The degree of decoupling is determined by the organization using the framework, not by the fact that there currently exist two different NIST frameworks.

There already appears to be an overlap between the two frameworks with the Identify and Protect Functions (see Figure 14). The Functions that do not overlap include Govern, Control, and Communicate for the privacy framework and Detect, Respond, and Recover for the cybersecurity framework. These appear to all be Functions that could assist organizations in both the privacy and security realms.

Function Unique Identifier	Function	Category Unique Identifier	Category
ID-P	Identify-P	ID.IM-P	Inventory and Mapping
		ID.BE-P	Business Environment
		ID.RA-P	Risk Assessment
		ID.DE-P	Data Processing Ecosystem Risk Management
GV-P	Govern-P	GV.PO-P	Governance Policies, Processes, and Procedures
		GV.RM-P	Risk Management Strategy
		GV.AT-P	Awareness and Training
		GV.MT-P	Monitoring and Review
CT-P	Control-P	CT.PO-P	Data Processing Policies, Processes, and Procedures
		CT.DM-P	Data Processing Management
		CT.DP-P	Disassociated Processing
CM-P	Communicate-P	CM.PO-P	Communication Policies, Processes, and Procedures
		CM.AW-P	Data Processing Awareness
PR-P	Protect-P	PR.PO-P	Data Protection Policies, Processes, and Procedures
		PR.AC-P	Identity Management, Authentication, and Access Control
		PR.DS-P	Data Security
		PR.MA-P	Maintenance
		PR.PT-P	Protective Technology
DE	Detect	DE.AE	Anomalies and Events
		DE.CM	Security Continuous Monitoring
		DE.DP	Detection Processes
RS	Respond	RS.RP	Response Planning
		RS.CO	Communications
		RS.AN	Analysis
		RS.MI	Mitigation
		RS.IM	Improvements
RC	Recover	RC.RP	Recovery Planning
		RC.IM	Improvements
		RC.CO	Communications

Figure 14. Overlap between the Cybersecurity and Privacy Frameworks. Source: [1].

Privacy engineers need to be able to respond to privacy breaches in a similar fashion to how organizations respond to security breaches. Sometimes these two types of breaches may be the result of the same event; therefore, it would make sense if the same framework were to be used when implementing privacy and security controls. Even the privacy framework Functions of Govern, Control, and Communicate would be beneficial to the security posture of a system. The security profile of an organization should be planned, deliberate, and a topic of discussion at the executive/c-suite level in the same manner as

privacy. Security, in the same manner as privacy, should not be isolated to only the technical advisors of an organization since the upper-level management aids in setting the risk tolerance of the organization, and this risk tolerance plays into the security profile as well. This reinforces why the cybersecurity of an organization could also benefit from applying some of the principles of the privacy framework.

While the integration of the two frameworks would be beneficial in terms leveraging more resources to tackle security and privacy issues that are largely interrelated, it would also add more simplicity to the federal development process since it would be easier to more seamlessly integrate one existing framework into an already established RMF process. When NIST SP 800-53 rev5 was released in September 2020, the main overall revision was that more emphasis was added to privacy specific controls [6]. NIST did not create a new Special Publication just for privacy controls and keep SP 800–53 for security-only controls. Whether or not this was just the easier solution, it still reinforces that privacy and security controls belong in the same space since there is overlap between the two.

NIST has been integrating the Cybersecurity Framework Functions and Categories into SP 800–53 rev5 as part of the controls. It may be only a matter of time before NIST begins integrating Functions and Categories of the privacy framework as well. It would certainly simplify the process for developers and engineers to be able to utilize one NIST framework and one NIST Special Publication that contains what is needed for both privacy and security risk assessments.

In the same vein as SP 800–53 rev5 including references to the NIST Cybersecurity Framework, SP 800–37 rev2 (“Risk Management Framework for Information Systems and Organizations”) released in December 2018 includes how the cybersecurity framework can be aligned with the RMF. Under Executive Order 13800, it is mandatory that federal agencies use the cybersecurity framework in conjunction with the RMF process [19]. The RMF Special Publication was updated before the NIST Privacy Framework was released in 2020, but it likely to only be a matter of time before the privacy framework is also required by executive order and SP 800–37 gets updated to include references to the NIST Privacy Framework. If the two frameworks were to be combined, it would allow for a

smoother integration into existing processes, and it would likely enhance the development of secure and private systems.

It is evident that NIST and the rest of the U.S. government have decoupled security and privacy from one another to some degree, but that does not seem to be fully intentional. It seems that cybersecurity was an issue that was tackled first, and then NIST began to work on privacy in a somewhat separate time and space. This is understandable as these represent relatively emerging technological issues; however, when the dust settles, it will become clearer that in this case less is more and that the frameworks would be consolidated into one "Risk Assessment Framework for Security and Privacy." Fundamentally, this reflects the idea that security and privacy are interrelated but still allows the distinction between the two concepts to be seen clearly based on developers' risk assessments of the application of the system being designed. From a practical standpoint, this reflects a better consolidation and integration of a single framework into already established federal documentation on risk management, security, and privacy.

B. THE CHECKLIST NATURE OF THE FRAMEWORK

Another aspect of the framework that warrants more clarification and discussion is the checklist nature of building target profiles using the framework. The privacy framework is set up in such a manner that a team of designers could scroll down the list of Functions, Categories, and Subcategories—checking off each row of the spread sheet after they have visited it and assessed that line item's value to their organization. This is not the ideal way to use the framework, nor was it the intended way that NIST envisioned that the framework would be used. Ideally, an organization has taken the responsibility to determine their own risk tolerance outside of applying the framework and make its privacy assessments based on the unique implementation of whatever system the organization is implementing (as mentioned in Chapter IV on risk modeling). In this case, the organization would use the privacy framework as a guide to further open up the privacy conversation and help identify any privacy issues that might have slipped through the cracks.

One of the problems with having the NIST Privacy Framework set up in a checklist fashion is that organizations might see this as a way to remove the responsibility of experienced privacy engineers. If one were to complete the privacy framework "checklist-

style," then they would assume that they have identified any and all privacy issues related to the system. While the privacy framework is adept at identifying certain privacy issues indirectly, it is incorrect to assume that every privacy issue becomes transparent or enters the privacy conversation based on an application of the framework. The NIST Privacy Framework does not claim to provide a roadmap to perfect or even near-perfect privacy; it only serves to help organizations take privacy into account when designing systems by promoting communication about privacy and encouraging collaboration [1].

This checklist type of thinking is problematic because of the prevalence of checklists in federal organizations—especially the military. The military gravitates toward checklists. Aviation fields use checklists for a wide variety of tasks. The assumption is that if the checklist is complete and all parameterized values are within the acceptable limits, then the task either has been or will be completed successfully. This mechanism obviously differs from the NIST Privacy Framework, but to what extent? The inexperienced practitioner would like a checklist that they would apply to the privacy space, and the results of that checklist would output whether a system is within the acceptable privacy limits. Unfortunately, privacy risk and privacy compliance are usually not that precise and hardly ever results in a binary outcome.

This becomes even more problematic when the desire for a military-style checklist manifests itself in the form of a compliance checklist. A pilot cannot take off until he or she has completed the take-off checklist successfully. It would not be too hard to imagine organizations in the federal government making the NIST Privacy Framework or some kind of equivalent framework mandatory for privacy compliance. This could quickly devolve into engineers and developers applying the framework like a checklist rather than a mechanism for awareness and discussion, and the results could be disastrous. It could also engender a practice of gaming the checklist, or even result in a proliferation of requests for waivers for completing the checklist or addressing specific items on the checklist.

In order to ensure that the NIST Privacy Framework remains just what it is, a framework and not a checklist, there are several steps that should be taken. The first step is to ensure that the NIST Privacy Framework remains voluntary for all organizations. The voluntary nature of the framework encourages analysis and discussion of privacy issues

whereas an involuntary application of the framework would likely lead to checklist type compliance.

The next step would be for NIST to emphasize the non-parameterized nature of privacy. They need to ensure it is clear that different people and organizations will apply the framework over a variety of situations which makes this different than a pilot applying a preflight checklist to an airframe where the outputs are easily calculated and quantified. Privacy compliance with the NIST Privacy Framework depends completely on an organization's dedication to ensure privacy to the maximum extent possible without overly impeding the organization's ability to fulfil it intended purpose. The privacy framework needs to clarify that privacy engineering is not an exact science or a binary problem. Like all forms of engineering, it is concerned with building systems and apps to engineering-tolerance and making engineering tradeoff decisions.

C. THE FRAMEWORK AS A RISK IDENTIFICATION TOOL

While the NIST Privacy Framework is a valuable resource for system engineers and developers, its usefulness lies in its support for identifying certain broad privacy issues, raising awareness about privacy issues, and offering guidance on implementing controls to manage privacy risks through risk-mitigation measures. The privacy framework admittedly was made to be flexible to accommodate different types of technologies and organizations, bringing into question the efficacy of applying the framework. An overarching question would be whether someone can use the framework to identify privacy issues for organizations and translate these issues into impacts or what needs to be done to fix them. The following sections examine questions that the privacy framework should strive to answer. These questions represent some of the core aspects of the framework, and as such, they provide valuable insight into the effectiveness of the framework as revealed by different case studies, to include this one on the DON Contact Tracing System.

1. Does it identify failures of the worst kind?

The ultimate goal of any risk mitigation framework is to identify the potential failures of the worst kind. Catastrophic failures that result in privacy breaches (and most times security breaches) affect companies' bottom lines, and in the case of the military,

they affect the ability to carry out the mission. An example would be a flaw in the design of a system that allows a certain unauthorized user to view customers' private data, which would be an issue, but this issue's impact could be magnified exponentially if this unauthorized user has the access where they can digitally download all of this private data and export it to third parties. Now a flaw that allowed one unauthorized user to view private data has been turned into a more severe privacy breach—the mass exfiltration of private data.

Could the use of the NIST Privacy Framework have resulted identifying that such a failure could occur? The NIST Privacy Framework does identify access controls as a form of privacy mechanism in the "Identity Management, Authentication and Access Control (PR.AC-P)" Category under the Protect (PR-P) Function; however, just because the framework identifies access controls as a privacy control, it provides little insight into how access controls should be implemented in order to the avoid the most catastrophic privacy failures of a system [1]. This aligns with the claim that the framework is in the simplest terms an identification tool; that is, the framework serves as guidance, but there is a considerable amount of work that needs to be done by the organization to define the impacts for privacy issues identified when applying the framework. Hiller and Russell share this concern in [54] by pointing out that the privacy framework assumes an organization will do all the work of determining impacts and likelihood before applying the framework.

A well-known modern-day example of a privacy breach that is not related to security was the Facebook-Cambridge Analytica data scandal. The British political consulting firm Cambridge Analytica influenced the 2016 U.S. election using personality profiles on Americans. The profiles were built using Facebook user data. The issue was that the data was obtained by Cambridge Analytica in what some would classify as a nefarious manner as many of the users unknowingly disclosed data about themselves and their Facebook friends. Since this event was strongly tied to a U.S. election, the unveiling of the creation and use of those personality profiles was instrumental in bringing digital privacy into the consciousness of the layperson and exposing privacy issues faced by Big Tech [55].

An article from Wired magazine claims that the CEO of Facebook was aware in 2012 of third-party apps collecting data on Facebook users' unwitting friends, but he did not see this as a significant risk to the company [55]. Facebook as well as the rest of the world now knows that it was a huge risk—one could say it was an unwarranted risk, both for the company and the users of Facebook. The point of this example is to illustrate that the NIST Privacy Framework is not complete in its guidance on identifying the severity of risk. Facebook's CEO saw a risk but did not foresee a huge impact over the horizon. Hopefully, use of the privacy framework would have helped to identify a risk like this one, but it is hard to imagine that the framework would have given any insight into the impact of the risk. That is the responsibility of the executives, engineers, and developers applying the framework. In addition, it is one thing to identify a risk, but it is another to do a good job of characterizing the risk in terms of severity and probability of occurrence.

The answer to the question of whether or not the NIST Privacy Framework provides adequate guidance for identifying all possible privacy failures at all levels of severity is *no*. NIST does not claim that its privacy framework will assist the user of the framework in determining the severity of a privacy risk; there is no silver bullet or one-size-fits-all solution, in either in privacy or security engineering. That aspect is left up to the executives, engineers, and developers. It should be made clearer that since the framework is more of an identification tool, it does little to illuminate the impacts of either privacy failures or privacy controls.

2. How does the framework lead from a policy control to a technical control?

The framework does not directly address how an established privacy policy can lead to technical controls. If the NIST Privacy Framework is going to provide value as a risk assessment tool, it needs to provide guidance on the appropriate technical controls being implemented. The DOD, DON, and DHA have privacy offices that implement privacy policies. In theory, the major tenets of these privacy policies should be able to be input into risk assessment tools that are based on the NIST Privacy Framework. Those tools should provide as output recommendations for technical controls that satisfy these policies (see Figure 15). While this is a simplified, black box scenario, the principle still holds. The question is whether the NIST Privacy Framework accomplishes this task.

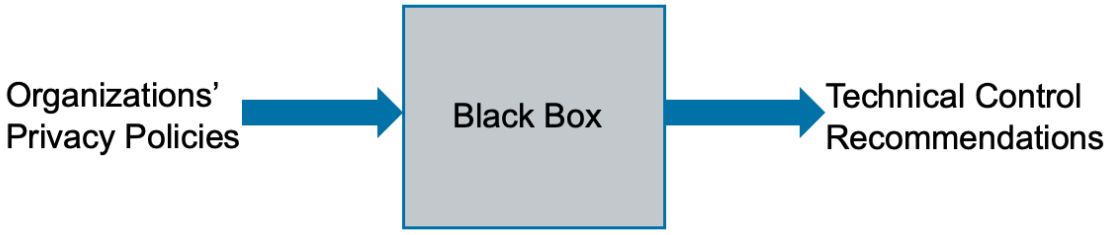

Figure 15. Desired Outcome of Privacy Frameworks and Risk Assessment Tools.

The DOD Instruction 5400.11 "DOD Privacy and Civil Liberties Programs" would be an appropriate case study for this question [23]. One of the Fair Information Practice Principles (FIPPs) that the DOD considers when evaluating information systems is minimization. While this is an open-ended tenet of most privacy policies, it is necessary for all DOD information systems. This means that the NIST Privacy Framework should recommend technical controls that implement this policy aspect.

If one were to scan through the NIST Privacy Framework, he or she would see the Data Processing Management Category under the Control Function. This category offers NIST SP 800-53 rev5 controls like event logging, monitoring, access control, and audit record review to aid in minimization of data in accordance with an organization's risk strategy. This represents the current granularity of the technical controls provided by the NIST Privacy Framework. It is the responsibility of the engineers and developers to apply these NIST SP 800-53 rev5 controls to the system being developed.

To answer the question whether the NIST Privacy Framework provides a link from policy to technical controls revolves around how specific engineers and developers expect these technical controls to be. Being a mechanism for identification, the framework will identify some broad technical controls that should accomplish the task that reflects the language in the privacy policy. One of the downsides to the framework is that it is currently only mapped to NIST SP 800-53 rev5 controls. The cybersecurity framework maps to controls from several different organizations in addition to the NIST SP 800-53 rev5. When more mappings have been added, the privacy framework will be more capable of translating policy controls to technical controls, but it is still important to point out that

these technical controls are useless without trained and experienced engineers and developers to implement them based on a sound risk strategy by their organization.

3. The system is not 100% private. Here are the risks. Do you accept the risks?

It has been established that there is no such thing as 100% privacy (or 100% security for that matter); privacy is assured through assessing the threats, the likelihood of those threats (i.e., frequency), the impact (i.e., severity) of those threats, and applying a risk determination based on the controls implemented to mitigate those threats. It is also a matter of there being an exploitable associated vulnerability that the threat actor can employ. This is the idea behind this final question about accepting privacy risks and the extent to which the NIST Privacy Framework exposes these risks and allows leadership to make informed decisions about the implementation of an information system like the DON Contact Tracing System.

As stated earlier, the privacy framework can help identify issues relating to privacy; it cannot determine likelihood or impact. Those are determined by each organization regarding each of their information systems. The privacy framework may be used to determine that the organization has an inadequate risk strategy, but it is ultimately up to that organization to reconcile that strategy with the needs of the organization and/or consumers. For example, Subcategory Identify, Risk Assessment (ID.RA-P) maps to NIST SP 800-53 rev5 controls for Risk Assessment (RA-3) and Privacy Impact Assessments (RA-8). The framework exists to "remind" organizations to ensure these controls exist and are adequate, but the control guide cannot provide extensive details on how these controls can be the most effective since each organization has different needs and resources. The framework cannot determine impact; it can only ensure the organization consider the impact and make the determination themselves.

Unlike impact or likelihood, the actual risks can be discovered by applying the privacy framework and implementing some of the privacy/security controls that are recommended. The privacy framework, under Identify, Data Processing Ecosystem Risk Management (ID.DE-P), maps to control SA-11 which is Developer Testing and Evaluation. One of the key components of ensuring the trustworthiness of a system (in

terms of either privacy or security) is to have a robust and extensive testing and evaluation program. This control could either identify that a current testing program is not adequate in terms of testing for privacy, or it could reveal other privacy issues during testing that need to be dealt with further.

Moreover, the answer to the question is that *it depends*. The NIST Privacy Framework cannot be used to directly obtain a full risk determination by the organization using it. Although there may be a temptation to use the framework as checklist as a means of compliance, it is not advisable to do so. The framework can be used to identify certain risks, but since it cannot directly determine impact or likelihood, it misses some of the pieces of the risk equation. This is to be expected, though. The authors of the framework emphasize the need for privacy expertise when using this framework. It does not claim that the privacy framework is the only thing an organization needs to ensure privacy, but they should still emphasize the existence of these restraints when applying the framework.

VII. CONCLUSION AND FUTURE WORK

Digital privacy research in the federal government, and specifically the DOD, is particularly important because unauthorized disclosure of privacy-sensitive data presents risks to accomplishing the enterprise's mission. Federal government civilian employee, military member, and defense contractor data is not explicitly tied to commercial endeavors as may be the case in the private sector. In the case of the federal government, this data takes on a different value when that data is tied to national security, force protection, and ongoing or future DOD operations.

The DON Contact Tracing System, which is being developed to curb the spread of COVID-19 and analyze social distancing measures, has the potential to collect private data on members of the DON workforce that adversaries could use to exploit and disrupt military operations and decision making [7]. Contact tracing data at a minimum includes time and proximity measurements between individuals, and this could lead to the reconstruction of organizational charts of a command or to the development of patterns of life of employees—both of which are serious OPSEC concerns. This information can also translate to adversaries making determinations about fleet readiness and future operations. Beyond the larger OPSEC concerns, leakage of privacy data can lead to coercion attacks against employees by those that wish to do harm to those individuals and exploit those them for the purposes of espionage and sabotage. As a bottom line, we observe that the privacy of DON personnel data must be considered a force protection issue with first-order implications for the design and operation of DON systems, not merely a legal hurdle to be cleared in a compliance-oriented manner.

Any system that collects private data is inherently vulnerable to data leaks. In practice, there is no such thing as perfect privacy in the same way there is no such thing as perfect security. If user privacy is jeopardized by the contact tracing system, then why implement a contact tracing system at all? The answer to that involves analyzing a utility versus privacy tradeoff. The contact tracing system is a solution toward stopping the spread of a disease, but does the privacy cost outweigh the utility of the contact tracing system [28]? Using a risk-management framework, one can perform a risk assessment of this

tension and other risk-related tradeoffs. This thesis contributes to conducting privacy risk assessments of this type.

Coincidentally, the outbreak of COVID-19 occurred about the same time as the release of the NIST Privacy Framework. The framework is meant to be used for conducting privacy risk assessments by facilitating a dialogue among stakeholders at all organizational levels about how to approach privacy when designing and implementing systems and applications [1]. The NIST Privacy Framework can contribute to the utility-versus-privacy calculation by providing the DON with a means to identify privacy risks in the contact tracing system.

To test this claim, we constructed a privacy-threat model of the contact tracing system. The model highlights the threats along the data flow of user data from collection by wearable devices to ingestion into Jupiter, the DON's enterprise data management system. The Jupiter system produces high-level decision-making products to leadership. After threat modeling, we applied the NIST Privacy Framework to the contact tracing system and focused on how the framework can be tailored toward identifying and addressing the same threats that we identified in the privacy threat model. This gave us a grasp on the extent to which the NIST Privacy Framework can be used in making risk assessments about privacy.

Following the application of the NIST Privacy Framework to the DON Contact Tracing System, we can divide our results into the following two categories: an assessment of and recommendations for the contact tracing system in terms of a privacy risk profile, and an assessment and recommendations for applying and improving the NIST Privacy Framework within DON.

One of our recommendations for engineers and developers involved in setting requirements for, implementing, and testing the DON contact tracing system or a similar system is that they be attentive to the intersection of policy, disassociation techniques, and security controls. All three areas need to be incorporated into the design of the system. Controls stem from policy, and if the privacy policy does not appropriately incorporate an organizational risk assessment, then the controls may not be effective.

The policy needs to be detailed at a technical level—not just in regard to legal compliance. One of the main results of threat modeling and applying the privacy framework is that more detail and consideration should be paid toward disassociation technologies and the policies governing them. In Chapter V, we reference the standards implemented by the GDRP and HIPAA concerning disassociation of private data. Specific technical requirements must flow down from these standards based on the risk tolerance of the DON for this specific application. Our research shows that the DON needs to look deeper into disassociation technologies. In addition, the DON needs to institute policies establishing its risk tolerance regarding de-identifying user data. We also recommend that the DON be wary of commercial-off-the-shelf (COTS) products for which the vendors make guarantees of privacy based on simplicity measures such as removing names from data sets. Disassociation of user data involves a much more technical and thorough risk-based approach.

We recommend that the DON Contact Tracing System adopt HIPAA's Expert Determination method, or something similar, to de-identify data [53]. We also recommend that the contact tracing data be collected and initially stored by BUMED personnel to ensure the chain of custody for PHI resides under a covered entity with an established standard for disassociation of user data. BUMED can then apply its de-identification standard before releasing data sets to Jupiter for use by the larger naval enterprise. Streamlining the data path through BUMED corresponds to applying the well-established principle of least privilege and supports privacy protection through robust data security.

We describe in this thesis how security is required for privacy, but it is not all that is required. Ensuring encryption of and controlled access to data supports confidentiality, but it is not sufficient for ensuring privacy. We make it clear that re-identifying users based on their collected data may have nothing to do with security if analysts have obtained the data sets through authorized channels. This is why we recommend that in addition to security testing by organizations like NSWC Crane, the DON should also develop a similar privacy testing program that includes privacy testing when assessing federal information systems. This would involve actions like applying a framework like the NIST Privacy Framework or assessing the likelihood of machine learning algorithms being able to re-identify datasets based on the addition or removal of quasi-identifiers.

Finally, we make an assessment of the NIST Privacy Framework and provide recommendations for improving it or applying it in the future. It became clear after applying the NIST Privacy Framework that the framework is primarily a guide for identifying privacy threats, and it should be used as such. As a guide for risk identification, the framework only addresses part of risk-management equation. The other important factors in risk management include assessed impact, assessed likelihood of occurrence, and an organization's risk tolerance [3]. The privacy framework is also intentionally vague and meant to be flexible. This leaves a lot of interpretation up to the organization applying it. This thesis asserts that the NIST Privacy Framework is not a complete solution for privacy-risk assessment.

The NIST Privacy Framework should not be used in a checklist manner for identifying threats and implementing controls. Doing so could be troublesome when determining compliance because the line items of the framework are not parameterized in a manner that readily supports binary decision-making on privacy. For this reason, we do not recommend that the NIST Privacy Framework be made mandatory as a compliance checker because privacy assurance involves levels of risk.

We recommend that NIST combine its cybersecurity and privacy frameworks into a single framework. Privacy and cybersecurity are separate but related concepts. Many of the control families overlap the two concepts. Combining the two frameworks would not only streamline the assessment process, but it would also better align with current standard publications that provide both privacy and security controls (i.e., NIST SP 800-53 rev5). Federal employees are already required to integrate the cybersecurity framework when assessing information systems, and it is likely only a matter of time before the referencing the privacy framework becomes a requirement.

The privacy and cybersecurity frameworks are similar in structure and remain insufficiently distinguishable from each other in several aspects. This work makes the argument that the privacy framework does not add enough value to privacy risk assessments to require it to be a separate guide. A single framework would be able to consolidate controls and mechanisms that apply to both security and privacy, but at the

same time, it would also be able to isolate and identify privacy-specific controls for use in a privacy risk assessment.

The NIST Privacy Framework is still a useful guide for privacy risk assessment in terms of identifying risk and identifying possible controls to those risks, but NIST needs to further the investigation of what mechanisms are needed for making determinations about the impact and likelihood of threats occurring. Additionally, the vagueness of the framework requires additional effort on the part of practitioners applying the framework in order to mold the framework around their intended privacy goals.

As for the DON Contact Tracing System, the privacy risk assessment depends on the risk tolerance of the DON at any point in time. Fortunately for the DON, guidance like that provided in the NIST Privacy Framework can assist in assessing privacy risk. The DON needs to ensure that it is also doing its due diligence to define its risk tolerance and adjust its technical and policy controls as laid out in this thesis's recommendations.

A. FUTURE WORK

The following sections present potential areas of research that would advance our understanding of privacy from two perspectives: technical mechanisms and policy mandates (including law). The first two areas of future work involve modeling and applying machine learning techniques to data sets that contain PII which might involve applications of differential privacy. The second two areas focus on investigating privacy governance surrounding the Jupiter system and how recently released privacy frameworks compare to each other.

1. Real-World Testing

Future work should focus first on testing the contact tracing system in a real-world application. The DON planned on using the USNA as a test bed for the technology, but just because the United States appears to be moving past COVID-19, one should not wait to continue this research until the next pandemic occurs. Testing the DON contact tracing system now will provide insight for its improvement and how to conduct privacy assessment of other similar Bluetooth systems for use in tracking personnel in the fleet. For example, a Naval vessel might want to use Bluetooth beacons for access control

(similar to RFID tags) or for monitoring foot traffic in certain spaces onboard a ship. The testing procedures would be similar for such scenarios, and the goal would be to measure how much private data is leaked about users' movements and what that leaked information could be used for (e.g., creation of high-fidelity social network graphs).

Ideally, the test would require DON employees and servicemembers of varying rank to wear Bluetooth tokens. The tester would need access to the identified and deidentified contact tracing data, and the tester could then apply for instance supervised machine learning techniques to the data in order to determine the difficultly of extracting private data for the wearers. The tester would test to determine the extent to which they can recover an organizational chart for the command or infer personal schedules that can lead to reidentifying those wearers; what we describe here is red-teaming for privacy. There is also the option of using synthetic data to test the privacy leakage of a Bluetooth Contact tracing system, but ground truth data would have to be constructed in order to use supervised machine learning techniques.

The starting point for this testing would be replicating existing analyses of risks to individual-level identification such as those in Eagle and Pentland or de Montjoye et al. [38], [45]. Similar techniques could be used or modified for evaluating the risk of individual-level inferences from data generated by a wearable contact tracing system.

2. Disassociation Technologies

In NIST's "Roadmap for Advancing the NIST Privacy Framework: A Tool for Improving Privacy through Enterprise Risk Management," NIST zeros in on several evolving areas of ongoing research with one of these being "De-identification Techniques and Re-identification Risks" [22]. This thesis also counts disassociation as a primary concern in the DON Contact Tracing System, especially after the contact tracing data has been transferred to BUMED and Jupiter. Further research is required to determine the effect that adding or removing quasi-identifiers to contact tracing data will have toward re-identification.

If the contact tracing data is available for testing, the best starting point would be to look at the works of de Montjoye, especially his work "Unique in the Crowd: The

privacy bounds of human mobility" [45]. In his works, he experiments with identifying unique patterns from user data from items such as mobile phones and credit cards. Future researchers could apply some of the techniques from his works to contact tracing data sets to determine the likelihood of re-identification. Quasi identifiers could also be added to the feature space in order to study how they improve that likelihood.

3. Jupiter System

This thesis presents a very high-level description of the Jupiter system, but further investigation is still needed into the intricacies and details of Jupiter and whether any undiscovered privacy or data governance issues are lingering beneath the surface. Jupiter is the consolidation of many subsystems and databases, and the relationships between these entities could reveal problem areas for privacy and data governance [12]. Of course, the designers of the system have sought approval for ATO through the appropriate channels, but there is room for further investigation of the data flow of data collected in the future as the ATO process does not explicitly consider privacy risk, which may evolve as the data within the system changes. This could reveal the details surrounding privacy implications in the technical, policy, and legal domains. A good starting point in this research would be to refer to the Kroll et al. paper about data governance [56].

An in-depth view of the privacy risk surrounding the use of Jupiter would entail work with the DON CIO Privacy Office. Some questions that could be answered are whether documentation that is more invasive and comprehensive than a PIA should be required for a system of this scope. The research presented in this thesis presents an argument that PIAs do not provide the most comprehensive risk assessment of a federal system. A combination of a privacy framework like the NIST Privacy Framework and a designated official or office that works directly for DON CIO that solely manages privacy aspects of this system may better serve privacy needs for this system. That would require a review of how DON CIO performs it privacy and cybersecurity functions regarding Jupiter.

4. The NIST Frameworks and the ISO/IEC 27701 Privacy Extension

The final area of future research would be to identify other federal systems dissimilar to the contact tracing system and applying the NIST Privacy Framework. The results of this research would also be focused on assessing the privacy framework itself and determine whether any further conclusions can be made about the effectiveness of the framework on other systems. This would provide evidence as to whether the framework can be applied to a wide spectrum of systems, and if this is not the case, what prevents its general applicability.

Furthermore, it would beneficial to compare the NIST Privacy Framework to the ISO/IEC 27701:2019 (the privacy extension to the ISO/IEC 27001 and ISO/IEC 27002) since both of these frameworks came out with a year of each other [57]. ISO appears to have made their privacy framework an extension of their security framework, whereas NIST released a privacy framework that is on the same level as its Cybersecurity framework. This thesis makes the recommendation of combining the NIST cybersecurity and privacy frameworks, and a comparison of the NIST frameworks and the ISO frameworks may provide some insight into whether it makes more sense to combine frameworks or keep them separate.

APPENDIX A. CONTACT TRACING SYSTEMS IN DEVELOPMENT AND USE

B. PAN-EUROPEAN PRIVACY PRESERVING PROXIMITY TRACING (PEPP-PT)

PEPP-PT is an organization incorporated in Switzerland that was created as a non-profit in March 2020 by a multi-national European team of scientists and technologists. The organization's goal is to facilitate the development of a digital pandemic management approach via a privacy-preserving, proximity-based contact tracing system [58]. PEPP-PT's main role is to provide standards, mechanisms, and services to countries and developers while fostering international cooperation, promoting knowledge sharing between organizations, and promoting the adoption of their proximity tracing systems [59]. PEPP-PT intends to aid the development of a proximity tracing system that is fully compliant with the General Data Protection Regulation (GDPR) and therefore interoperable when traveling between European countries [58], [60].

In addition to international interoperability, all the implementations of PEPP-PT must integrate with smartphone technology and local IT infrastructure. PEPP-PT requires the use of smartphone applications for contact tracing because it is a ubiquitous digital technology that can be retrofitted as a measuring device using Bluetooth-based proximity measurements and can communicate to users that are at risk of infection [59], [60].

This international initiative is the umbrella organization for several proximity tracing implementations that follow the interoperability standards and requirements set forth by PEPP-PT. These standards and requirements range from being secure and privacy-preserving to being approved by national health authorities all while being able to be dismantled when no longer needed [60].

The systems currently under the PEPP-PT project are detailed below and include DP-3T, NTK, and ROBERT. They all use Bluetooth Low Energy (BLE) for proximity measurements, but they differ primarily in terms of their centralized or decentralized approach to proximity tracing.

1. Decentralized Privacy-Preserving Proximity Tracing (DP-3T)

DP-3T is a proximity tracing system that uses "a smartphone app that continuously broadcasts an ephemeral, pseudo-random ID representing the user's phone and also records the pseudo-random identifiers observed from smartphones in close proximity" [58]. If a user is diagnosed with COVID-19, the user can upload to a backend server all the pseudo-random identifiers that were previously broadcast from their phone over a predesignated time interval. The backend server will then distribute anonymous exposure information to other users that they can cross reference with the pseudo-random identifiers that they collected over the same predesignated time interval. If the user checking this exposure information finds a match between a pseudo-random identifier of an infected user and a pseudo-random identifier that they have collected previously and stored in their smartphone, the potentially affected user can take the proper precautions by seeking medical attention for testing and potentially quarantining themselves [61].

Users uploading their representations of their ephemeral identifiers to the backend server upon diagnosis is voluntary, and users have to continuously query the backend server for updated exposure information [61].

The most important feature of this system is that it is decentralized, meaning the backend server does not perform any processing, and it only acts as a communication platform between users. A compromise of this server would not affect the privacy of users providing or receiving information to or from this backend server [61].

DP-3T proposes three different implementations that include low-cost decentralized proximity tracing, unlikable decentralized proximity tracing, and hybrid decentralized proximity tracing. For each implementation, the ephemeral identifiers are generated cryptographically using a random seed linked to a time value (day, epoch, time window) within each user's smartphone. Once generated, these ephemeral identifiers are broadcast to other smartphones, and the receiving smartphones store a version of the ephemeral identifier collected, the exposure measurement, and the day or time window when the ephemeral identifier was received [61].

If a user is diagnosed with COVID-19, the user can upload to the backend server the random seed used to generate the ephemeral IDs as well as the time value linked to that

seed. This is the information the backend server broadcasts to the other users, and the other users use these two pieces of information to regenerate the ephemeral identifiers and calculate the risk of infection by comparing them with the stored ephemeral identifiers on their smartphones [61].

The three implementation designs mainly differ in regard to bandwidth and storage requirements, but each design has built-in mechanisms to protect privacy. All implementations allow the user to upload a *representation* of the ephemeral identifiers (random seeds and linked time value) to the backend server rather than the raw ephemeral identifiers. The random seed used to generate ephemeral IDs will also change at predetermined time intervals, including when the seed is uploaded to the backend server, which makes it more difficult for a network observer to identify users based on ephemeral identifiers. Some of the implementations also broadcast dummy traffic to the backend server to further deter and mislead network observers [61].

While DP-3T claims to be privacy preserving, some aspects of the system have been identified as inherently bad for privacy. In addition to a slew of attacks by an adversary that include false alerts, false reports, replay attacks, and relay attacks; DP-3T provides a new vector to track people using the smartphone app. Vaudenay claims that Bluetooth broadcasting, at the very least, will reveal the presence of a smartphone using the DP-3T app [29]. Some of the more extreme threats to privacy could include deanonymizing users as well as adversaries coercing users to give up their personal data which is all stored on their smartphones in accordance with the DP-3T system. The DP-3T system does collect a minimal amount of information from users, but this information can reveal a lot if analyzed by an adversary [29].

2. NTK and ROBERT

PEPP-PT NTK and PEPP-PT ROBERT are similar systems, implemented in Germany and France respectively, and they differ from the DP-3T mainly by being characterized as centralized systems [62]. This means that infected users send received ephemeral identifiers to a central server rather that user's own broadcasted ephemeral identifiers. The central server then performs all the risk calculations rather than those calculations being performed locally on the user's smartphone. The central server will use

these calculations to determine which users are at risk and then notify those users via the smartphone app [63]–[65].

The creation of ephemeral identifiers occurs in a similar manner as with DP-3T, but with NTK/ROBERT, the ephemeral identifiers are derived at the backend server and sent to users. This allows the NTK/ROBERT server to keep a copy of all the recent ephemeral identifiers created, and they can be used for risk calculations as infected users report received ephemeral identifiers to the backend server [63], [64].

Since the backend server derives all the ephemeral identifiers from a permeant identifier that users will register with the system, the backend server can decrypt any ephemeral identifier back to that permanent identifier. This is one of the primary privacy concerns with NTK/ROBERT and centralized systems in general. There is always the fear that this kind of system can be used for mass surveillance of citizens as it would not be too difficult to deanonymize a user by linking their permanent identifier to other identifiers external to the system [63], [64].

This centralized system also creates an environment where the backend server can construct a social graph of users using proximity values and timestamps received by users. The more people that report infection to the backend server, the better the backend server is able to expand on and fine tune the social graph of users. This means that users not at risk could have a portion of their social graphs exposed by the contact submissions of others [63], [64].

C. PRIVATE AUTOMATED CONTACT TRACING (PACT)

PACT, much like PEPP-PT, is a collaboration among several universities, private research centers, and development centers that are working toward designing an exposure detection system using personal digital communication devices. The collaboration is led by MIT Computer Science and Artificial Intelligence Laboratory (CSAIL), MIT Internet Policy Research Initiative, Massachusetts General Hospital Center for Global Health, and MIT Lincoln Laboratory [66]. PACT is similar to the DP-3T system in many regards; the most prominent similarity is that they both use a Bluetooth Low Energy, decentralized

approach for proximity tracing, and both organizations agree that privacy is a paramount concern in the development process [66], [67].

PACT enabled devices broadcast "chirps" that are created in the same manner as ephemeral identifiers in the DP-3T system; the only slight difference is that PACT is modeling its proposal on Apple's Find My protocol [67]. Whether this was intentional, this will allow PACT to more easily integrate with Apple devices when the application is implemented.

It is important to note that with this system, the user is given autonomy over the life cycle of the system. In addition to the user being able to make the choice whether or not to use the system, all risk calculations are determined on the user's device, and the user controls the process of uploading their broadcasted "chirp" information to the exposure database. Furthermore, the user also has the option of uploading only a portion of their data to the database or delay the upload of their data if they so choose [67].

As stated earlier, privacy controls are baked into the system. PACT claims it "provides the highest amount of privacy possible for an automated contact tracing system, even against a technologically sophisticated attacker with the ability to eavesdrop on all nearby BLE transmissions and write custom software that interacts with all devices in close proximity" [67]. Nevertheless, PACT does acknowledge that privacy in the system is not absolute, and that there are privacy flaws inherent to any automated contact tracing system. This is why the PACT collaboration emphasizes that users are still expected to consider the tradeoffs between social health and civil liberties [67].

D. APPLE / GOOGLE

One of the lines of efforts of the PACT collaboration described above is to assess the exposure notification software developed and distributed by Apple and Google since Apple and Google announced a joint effort in April 2020 to develop a system using Bluetooth on their mobile devices to reduce the spread of COVID-19 [66], [68]. The Apple and Google system is designed to be rolled out in two phases. In the first phases that started in May 2020, Apple and Google allowed contact tracing apps developed by public health authorities to work across their devices using APIs that Apple and Google made available.

In the second phase, Apple and Google are updating the operating systems of their devices to send out and listen for Bluetooth beacons without requiring the app to be installed. If a user is diagnosed to be at risk, they will be prompted to download an official app [68].

The Apple and Google system resembles the DP-3T system in many ways. The system is decentralized; therefore, all risk calculations take place on the user's device. The ephemeral identifiers are generated in a similar manner to DP-3T where Apple and Google use rolling identifiers to prevent linking and wireless tracking. When a user is diagnosed as positive for COVID-19, they have the choice to upload their rolling identifiers that they broadcast during a specified time interval [69].

In addition to the rolling identifiers to help prevent tracking, Apple and Google have built in other privacy enhancing features that have also been present in PACT and PEPP-PT requirements. Apple and Google stress the autonomy of the users in being able to decide whether they want to use the app, whether they want to share data, and additionally they can also control which data they want to share [68]. Apple and Google further stress that the technology will only share data with public health authorities. In congruence with PACT and PEPP-PT requirements, Apple and Google have also acknowledged that they will be able to disable exposure notifications when this technology's functionality is no longer needed [68].

E. TRACETOGETHER

TraceTogether is a contact tracing system developed by Singapore's Government Technology Agency and Ministry of Health, and it is considered the first national deployment of a Bluetooth-based contact tracing system in the world [33]. This system is the most similar to PEPP-PT NTK/ROBERT since it uses a centralized system to generate temporary identifiers and calculate risk assessments. A central server contains the secret key to encrypt and decrypt temporary identifiers; therefore, infected users must upload their encounter history (temporary identifiers received) to the central server in order for the health authorities in Singapore to calculate a risk assessment [33].

While TraceTogether appears to be implemented as a centralized system in Singapore, BlueTrace, the privacy preserving protocol that TraceTogether uses, claims that

the system is actually a hybrid model of contact tracing. This is based on encounter messages being stored in a decentralized peer-to-peer fashion between users. Only when someone is diagnosed with COVID-19 does the user need to interact with a central server [33]. This is a somewhat clever distinction made by BlueTrace, but it does not stop others from classifying it as a centralized system along with ROBERT and NTK [30]. BlueTrace also claims that TraceTogether could be implemented as a completely decentralized system, but they do not recommend using decentralized systems for contact tracing, citing it as an easy way to abuse the system by users self-reporting false alerts [33].

Much like the other systems previously discussed, TraceTogether considers privacy a priority which it manifests in aspects like rotating temporary identifiers to prevent tracking users, storing encounter histories locally on user devices, and giving user control over having their data stored with the health authority. The most concerning privacy aspect of the TraceTogether system is that users must register their phone number with the health authority in order to use the app. This gives the health authority (central server) a piece of personally identifiable information to link to the users. Because of this, TraceTogether cannot avoid the same concerns as other centralized systems like ROBERT/NTK where a high level of trust must be placed on the government and health entities that have access to that central server [33], [37], [70]. Singapore prefers this centralized system specifically due to the fact that it allows health authorities to reconstruct more of a social graph when contact tracing. For that same reason, Singapore declined to use Apple and Google's system [36].

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B. NIST PRIVACY FRAMEWORK APPLICATION

The goal of this Appendix is to operationalize the NIST Privacy Framework in the context of the DON Contact Tracing System. In other words, does the NIST Privacy Framework lead the DON to a point where it can accomplish its goals while preserving privacy in the process? Can the NIST Privacy Framework mappings to the NIST SP 800-53 rev5 controls sufficiently address the privacy concerns discussed in Chapter III? Is the NIST Privacy Framework able to accomplish what it actually claims when applied to this contact tracing system?

The above questions lay the groundwork for applying the NIST Privacy Framework to the contact tracing system in this chapter. This involves discussing the how the Core functions are mapped to the NIST SP 800-53 rev5, reiterating the goals of the contact tracing system, and providing criteria for control selection in this particular application. Controls are then chosen from each Core function if they are applicable to the goals and/or privacy concerns of the contact tracing system.

For this research, the NIST Privacy Framework is being applied after conducting the privacy threat model in Chapter IV. The threat model is meant to provide a foundation of the system and its privacy threats in order to streamline the application of the NIST Privacy Framework.

A. METHODOLOGY

The following section details control selection for the DON Contact Tracing System. This involves identifying a pool of controls that deal with security and privacy (NIST SP 800-53 rev5 in this case), evaluating the goals of the system, and the choosing criteria on which to base control selection.

1. Core and Mappings to NIST SP 800-53 rev5

Along with the release of the NIST Privacy Framework, NIST has also provided on its website an Excel spreadsheet that maps the Privacy Framework's Core Functions, Categories, and Subcategories to NIST SP 800-53 rev5 controls [71]. NIST makes the claim that "Organizations should not assume there is a one-to-one relationship between the

SP 800–53 controls and the Subcategories" [71]. While this makes sense since every system is unique and may require a different approach to achieving privacy goals, at some level these organizations need to apply controls to achieve its goals. If NIST is going to provide a mapping to controls, these controls should be evaluated along with the framework itself since these control recommendations may have implications for the overall effectiveness of the framework.

Since NIST has established NIST SP 800-53 rev5 controls as the most relevant baseline controls that map to the Privacy Framework, this annex will use the controls provided in the mapping to perform the assessment of the contact tracing system.

The first step in applying the NIST Privacy Framework is identifying Core Functions, Categories, and Subcategories that an organization wants to prioritize to help it manage privacy risk. This allows an organization to build a profile of its risk management needs which can then be evaluated via implementation tiers. The scope of this assessment will involve only choosing the Core Functions, Categories, Subcategories, and NIST SP 800-53 rev5 controls that are applicable to the contact tracing system. This essentially constructs a profile of the system, and it should also recommend controls for mitigating privacy concerns within that profile.

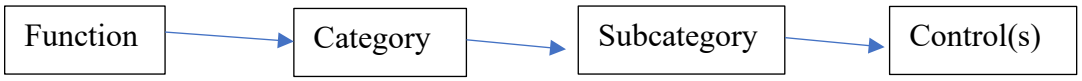

Figure 16. NIST Privacy Framework Organization.

2. Criteria for Control Selection

The Core Functions map to 228 different controls in NIST SP 800-53 rev5. It would be exhaustive to apply every control to the DON Contact Tracing System in this annex; therefore, this annex applies criteria for the selection of controls to evaluate the system against. These criteria are meant to include controls that apply specifically to the DON Contact tracing system and its somewhat unique privacy concerns. Controls that can be applied broadly to most systems being evaluated will be avoided in this assessment.

These broad categories of controls that will be omitted include controls that may focus on testing, auditing, maintenance, resilience, and supply chain. While these categories of control are important to the assessment of the contact tracing system, they represent categories of privacy/security controls that should be applied to any system procured or developed by the DON regardless of function or purpose. This mainly limits the control selection to technical privacy-specific controls as well as policy and procedure controls since privacy is very dependent on policy measures in a contact tracing system of this size.

3. Goals of the System and Privacy Concerns (Consolidated from Chapter III)

It is important to reiterate and/or consolidate the goals of the DON Contact Tracing system (shown in Table 6) in order to better understand the rationale behind which Core Functions, Categories, Subcategories, and controls are used. The goals of the system along with specific privacy concerns of the system will determine how controls are chosen.

Table 6. Goals of the Contact Tracing System and Privacy Concerns

<table>
<tr><td colspan="3">Goals of the Contact Tracing System</td></tr>
<tr><td colspan="3">The DON Contact Tracing System shall mitigate the spread of COVID-19 and measure the effectiveness of social distancing measures at Naval commands.
The system shall also protect user privacy in accordance with applicable laws, policies, and regulations.
The system shall protect Personally Identifiable Information that could lead to operational security infractions or data being leaked about mission readiness of a command.</td></tr>
<tr><td>Policy Concerns:</td><td colspan="2">Data governance between multiple organizations
Access control policies
Minimization of data</td></tr>
<tr><td rowspan="6">Technical Concerns:</td><td rowspan="4">Wearable Device Collection</td><td>Centralized versus Decentralized</td></tr>
<tr><td>Bluetooth versus GPS</td></tr>
<tr><td>Automated versus Human in the loop</td></tr>
<tr><td>Wearable versus Phone App</td></tr>
<tr><td rowspan="2">Distributed Architecture</td><td>Cloud Architecture</td></tr>
<tr><td>Priority of Data</td></tr>
<tr><td colspan="2">De-identification</td></tr>
</table>

The NIST Privacy Framework needs to at least address the above concerns over privacy in order for the DON Contact Tracing System to fulfill its goals.

4. Claims of the Framework

The other consideration to this assessment of the DON Contact Tracing System is having a realistic expectation of what the NIST Privacy Framework is able to provide in ensuring privacy in a system. NIST claims that the Privacy Framework's purpose is to "manage privacy risk" [1]. This includes assisting organizations with privacy risk

assessments, creating privacy mappings to informative references, strengthening accountability with stakeholders, applying privacy to the System Development Life Cycle (SDLC), establishing or improving privacy programs, assessing the data processing ecosystem, and informing buying decisions [1].

The DON's goals broadly align with these stated purposes of the framework; however, those claims of the framework are meant to be very broad and vague so that the framework can remain flexible to any technology, sector, law, or jurisdiction. If the NIST Privacy Framework is unable to provide the results or granularity that the DON is expecting, the DON should consider whether or not the NIST Privacy Framework is an acceptable framework for this task (see Chapter VI for more details on framework applicability).

B. RESULTS TABLE

The following table shows the chosen NIST SP 800-53 rev5 controls that were chosen based on the needs of the contact tracing system. These controls can and should be combined with other common security/privacy controls when designing the system in order to increase the privacy of the system. Based on the way the NIST Privacy Framework was designed, this would constitute the first step of applying the framework by choosing Core Functions, Categories, and Subcategories that are prioritized by the organization. The table below would represent a version of a Profile that the DON is trying to establish in order to assess privacy and make risk decisions when continuing to design and implement the system.

Table 7 (shown below) was taken directly from the NIST SP 800-53 rev5 mapping spreadsheet provided by NIST [71]. The only modifications to the original spreadsheet occur in the "NIST SP 800-53, Revision 5, Control" column. For this thesis, this column is separated into two columns that show the control identifier in NIST SP 800-53 rev5 and the associated control name for ease of reading. NIST SP 800-53 rev5 controls that were not applicable to the contact tracing system are omitted from the spreadsheet.

Table 7. NIST Privacy Framework and Chosen SP 800–53 Controls. Adapted from [1], [6], [71].

NIST Privacy Framework Core			NIST SP 800-53, Revision 5, Control	
Function	Category	Subcategory	Control ID	Control Name
IDENTIFY-P (ID-P): Develop the organizational understanding to manage privacy risk for individuals arising from data processing.	**Inventory and Mapping (ID.IM-P):** Data processing by systems, products, or services is understood and informs the management of privacy risk.	**ID.IM-P1:** Systems/products/services that process data are inventoried.	CM-12	Information Location
			CM-13	Data Action Mapping
		ID.IM-P2: Owners or operators (e.g., the organization or third parties such as service providers, partners, customers, and developers) and their roles with respect to the systems/products/services and components (e.g., internal or external) that process data are inventoried.	CM-13	Data Action Mapping
		ID.IM-P3: Categories of individuals (e.g., customers, employees or prospective employees, consumers) whose data are being processed are inventoried.	CM-13	Data Action Mapping
		ID.IM-P4: Data actions of the systems/products/services are inventoried.	CM-13	Data Action Mapping
		ID.IM-P5: The purposes for the data actions are inventoried.	CM-13	Data Action Mapping
			PT-2	Authority to Process PII
		ID.IM-P6: Data elements within the data actions are inventoried.	CM-13	Data Action Mapping

NIST Privacy Framework Core			NIST SP 800-53, Revision 5, Control	
Function	Category	Subcategory	Control ID	Control Name
		ID.IM-P7: The data processing environment is identified (e.g., geographic location, internal, cloud, third parties).	CM-12	Information Location
			CM-13	Data Action Mapping
		ID.IM-P8: Data processing is mapped, illustrating the data actions and associated data elements for systems/products/services, including components; roles of the component owners/operators; and interactions of individuals or third parties with the systems/products/services.	CM-13	Data Action Mapping
	Business Environment (ID.BE-P): The organization's mission, objectives, stakeholders, and activities are understood and prioritized; this information is used to inform privacy roles, responsibilities, and risk management decisions.	**ID.BE-P1:** The organization's role(s) in the data processing ecosystem are identified and communicated.		
		ID.BE-P2: Priorities for organizational mission, objectives, and activities are established and communicated.	PM-11	Mission Business Process Definition
		ID.BE-P3: Systems/products/services that support organizational priorities are identified and key requirements communicated.	RA-9	Criticality Analysis
	Risk Assessment (ID.RA-P): The organization	**ID.RA-P1:** Contextual factors related to the systems/products/services and	CM-13	Data Action Mapping

NIST Privacy Framework Core			NIST SP 800-53, Revision 5, Control	
Function	Category	Subcategory	Control ID	Control Name
	understands the privacy risks to individuals and how such privacy risks may create follow-on impacts on organizational operations, including mission, functions, other risk management priorities (e.g., compliance, financial), reputation, workforce, and culture.	the data actions are identified (e.g., individuals' demographics and privacy interests or perceptions, data sensitivity and/or types, visibility of data processing to individuals and third parties).	RA-3	Risk Assessment
			RA-8	Privacy Impact Assessment
		ID.RA-P2: Data analytic inputs and outputs are identified and evaluated for bias.		
		ID.RA-P3: Potential problematic data actions and associated problems are identified.	CM-13	Data Action Mapping
			RA-3	Risk Assessment
			RA-8	Privacy Impact Assessment
		ID.RA-P4: Problematic data actions, likelihoods, and impacts are used to determine and prioritize risk.	RA-3	Risk Assessment
			RA-8	Privacy Impact Assessment
		ID.RA-P5: Risk responses are identified, prioritized, and implemented.	RA-8	Privacy Impact Assessment
	Data Processing Ecosystem Risk Management (ID.DE-P): The organization's priorities, constraints, risk tolerance, and assumptions are	**ID.DE-P1:** Data processing ecosystem risk management policies, processes, and procedures are identified, established, assessed, managed, and agreed to by organizational stakeholders.	SA-9	External System Services
			SR-4	Provenance
			PM-9	Risk Management Strategy

NIST Privacy Framework Core			NIST SP 800-53, Revision 5, Control	
Function	Category	Subcategory	Control ID	Control Name
	established and used to support risk decisions associated with managing privacy risk and third parties within the data processing ecosystem. The organization has established and implemented the processes to identify, assess, and manage privacy risks within the data processing ecosystem.	**ID.DE-P2:** Data processing ecosystem parties (e.g., service providers, customers, partners, product manufacturers, application developers) are identified, prioritized, and assessed using a privacy risk assessment process.	RA-3	Risk Assessment
			RA-8	Privacy Impact Assessment
		ID.DE-P3: Contracts with data processing ecosystem parties are used to implement appropriate measures designed to meet the objectives of an organization's privacy program.	SA-9	External System Services
		ID.DE-P4: Interoperability frameworks or similar multi-party approaches are used to manage data processing ecosystem privacy risks.		
		ID.DE-P5: Data processing ecosystem parties are routinely assessed using audits, test results, or other forms of evaluations to confirm they are meeting their contractual, interoperability framework, or other obligations.	SA-9	External System Services

NIST Privacy Framework Core			NIST SP 800-53, Revision 5, Control	
Function	Category	Subcategory	Control ID	Control Name
GOVERN-P (GV-P): Develop and implement the organizational governance structure to enable an ongoing understanding of the organization's risk management priorities that are informed by privacy risk.	**Governance Policies, Processes, and Procedures (GV.PO-P):** The policies, processes, and procedures to manage and monitor the organization's regulatory, legal, risk, environmental, and operational requirements are understood and inform the management of privacy risk.	**GV.PO-P1:** Organizational privacy values and policies (e.g., conditions on data processing such as data uses or retention periods, individuals' prerogatives with respect to data processing) are established and communicated.	all -1 controls	Policy and Procedures
		GV.PO-P2: Processes to instill organizational privacy values within system/product/service development and operations are established and in place.	SA-3	System Development Life cycle
		GV.PO-P3: Roles and responsibilities for the workforce are established with respect to privacy.	all -1 controls	Policy and Procedures
			PM-18	Privacy Program Plan
			PM-19	Privacy Program Leadership Role
		GV.PO-P4: Privacy roles and responsibilities are coordinated and aligned with third-party stakeholders (e.g., service providers, customers, partners).	PM-18	Privacy Program Plan
		GV.PO-P5: Legal, regulatory, and contractual requirements regarding privacy are understood and managed.	all -1 controls	Policy and Procedures
			PM-3	Information Security and Privacy Resources
			PM-7	Enterprise Architecture

NIST Privacy Framework Core			NIST SP 800-53, Revision 5, Control	
Function	Category	Subcategory	Control ID	Control Name
		GV.PO-P6: Governance and risk management policies, processes, and procedures address privacy risks.	PM-18	Privacy Program Plan
			PM-23	Data Governance Body
			PM-28	Risk Framing
			RA-3	Risk Assessment
			RA-8	Privacy Impact Assessments
	Risk Management Strategy (GV.RM-P): The organization's priorities, constraints, risk tolerances, and assumptions are established and used to support operational risk decisions.	**GV.RM-P1:** Risk management processes are established, managed, and agreed to by organizational stakeholders.	PM-9	Risk Management Strategy
			PM-28	Risk Framing
		GV.RM-P2: Organizational risk tolerance is determined and clearly expressed.	PM-9	Risk Management Strategy
		GV.RM-P3: The organization's determination of risk tolerance is informed by its role(s) in the data processing ecosystem.	PM-28	Risk Framing
	Awareness and Training (GV.AT-P): The organization's workforce and third parties engaged in data processing are provided privacy awareness education and are trained to perform their privacy-related	**GV.AT-P1:** The workforce is informed and trained on its roles and responsibilities.		
		GV.AT-P2: Senior executives understand their roles and responsibilities.		
		GV.AT-P3: Privacy personnel understand their roles and responsibilities.		

NIST Privacy Framework Core			NIST SP 800-53, Revision 5, Control	
Function	Category	Subcategory	Control ID	Control Name
	duties and responsibilities consistent with related policies, processes, procedures, and agreements and organizational privacy values.	**GV.AT-P4:** Third parties (e.g., service providers, customers, partners) understand their roles and responsibilities.		
	Monitoring and Review (GV.MT-P): The policies, processes, and procedures for ongoing review of the organization's privacy posture are understood and inform the management of privacy risk.	**GV.MT-P1:** Privacy risk is re-evaluated on an ongoing basis and as key factors, including the organization's business environment (e.g., introduction of new technologies), governance (e.g., legal obligations, risk tolerance), data processing, and systems/products/services change.	CM-4	Impact Analyses
			RA-3	Risk Assessment
			RA-8	Privacy Impact Assessment
		GV.MT-P2: Privacy values, policies, and training are reviewed and any updates are communicated.	all -1 controls	Policy and Procedures
		GV.MT-P3: Policies, processes, and procedures for assessing compliance with legal requirements and privacy policies are established and in place.		
		GV.MT-P4: Policies, processes, and procedures for communicating progress on managing privacy risks are established and in place.		
			CM-4	Impact Analyses

NIST Privacy Framework Core			NIST SP 800-53, Revision 5, Control	
Function	Category	Subcategory	Control ID	Control Name
		GV.MT-P5: Policies, processes, and procedures are established and in place to receive, analyze, and respond to problematic data actions disclosed to the organization from internal and external sources (e.g., internal discovery, privacy researchers, professional events).	RA-3	Risk Assessment
			RA-8	Privacy Impact Assessments
			SI-19(8)	De-identification | Motivated Intruder
		GV.MT-P6: Policies, processes, and procedures incorporate lessons learned from problematic data actions.	all -1 controls	Policy and Procedures
		GV.MT-P7: Policies, processes, and procedures for receiving, tracking, and responding to complaints, concerns, and questions from individuals about organizational privacy practices are established and in place.		
CONTROL-P (CT-P): Develop and implement appropriate activities to enable organizations or individuals to manage data with sufficient granularity to manage privacy risks.	**Data Processing Policies, Processes, and Procedures (CT.PO-P):** Policies, processes, and procedures are maintained and used to manage data processing (e.g., purpose, scope,	**CT.PO-P1:** Policies, processes, and procedures for authorizing data processing (e.g., organizational decisions, individual consent), revoking authorizations, and maintaining authorizations are established and in place.	PT-1	Policy and Procedures
			PT-2	Authority to Process PII
			PT-3	PII Processing Purposes
			PT-4	Consent
			AC-1	Policy and Procedures

NIST Privacy Framework Core			NIST SP 800-53, Revision 5, Control	
Function	Category	Subcategory	Control ID	Control Name
	roles and responsibilities in the data processing ecosystem, and management commitment) consistent with the organization's risk strategy to protect individuals' privacy.	**CT.PO-P2:** Policies, processes, and procedures for enabling data review, transfer, sharing or disclosure, alteration, and deletion are established and in place (e.g., to maintain data quality, manage data retention).	AC-3(14)	Access Enforcement | Individual Access
			CM-9	Configuration Management Plan
		CT.PO-P3: Policies, processes, and procedures for enabling individuals' data processing preferences and requests are established and in place.	AC-1	Policy and Procedures
			AC-3(14)	Access Enforcement | Individual Access
			PT-1	Policy and Procedures
			PT-4	Consent
		CT.PO-P4: A data life cycle to manage data is aligned and implemented with the system development life cycle to manage systems.	PL-8	Security and Privacy Architectures
			SA-3	System Development Life Cycle
			SA-8	Security and Privacy Engineering Principles
			SA-10	Developer Configuration Management
			SA-15	Development Process, Standards, and Tools
			SA-17	Developer Security and Privacy Architecture and Design
	Data Processing Management (CT.DM-P): Data are managed consistent with the organization's risk strategy to protect individuals' privacy, increase manageability, and enable the	**CT.DM-P1:** Data elements can be accessed for review.	AC-2	Account Management
			AC-3	Access Enforcement
			AC-3(14)	Access Enforcement | Individual Access
		CT.DM-P2: Data elements can be accessed for transmission or disclosure.	AC-2	Account Management
			AC-3	Access Enforcement
			AC-4	Information Flow Enforcement
			AC-21	Information Sharing
			CM-6	Configuration Settings
			AC-2	Account Management

NIST Privacy Framework Core			NIST SP 800-53, Revision 5, Control	
Function	Category	Subcategory	Control ID	Control Name
	implementation of privacy principles (e.g., individual participation, data quality, data minimization).	**CT.DM-P3:** Data elements can be accessed for alteration.	AC-3	Access Enforcement
			CM-6	Configuration Settings
		CT.DM-P4: Data elements can be accessed for deletion.	AC-2	Account Management
			AC-3	Access Enforcement
			CM-6	Configuration Settings
		CT.DM-P5: Data are destroyed according to policy.	SI-12(3)	Information Management and Retention | Information Disposal
		CT.DM-P6: Data are transmitted using standardized formats.	SI-10	Information Input Validation
		CT.DM-P7: Mechanisms for transmitting processing permissions and related data values with data elements are established and in place.	AC-16	Security and Privacy Attributes
			PT-2(1)	Authority to Process PII | Data Tagging
			PT-3(1)	PII Processing Purposes | Data Tagging
			SC-7(24)	Boundary Protection | PII
			SI-18(2)	PII Quality Operations | Data Tags
		CT.DM-P8: Audit/log records are determined, documented, implemented, and reviewed in accordance with policy and incorporating the principle of data minimization.	AU-1	Policy and Procedures
			AU-13	Monitoring for Information Disclosure
			AU-16	Cross-Organizational Audit Logging
		CT.DM-P9: Technical measures implemented to manage data processing are tested and assessed.	CM-4(2)	Impact Analyses | Verification of Controls
			SC-16(1)	Transmission of Security and Privacy Attributes | Integrity Verification
			SI-19(8)	De-identification | Motivated Intruder

NIST Privacy Framework Core			NIST SP 800-53, Revision 5, Control	
Function	Category	Subcategory	Control ID	Control Name
		CT.DM-P10: Stakeholder privacy preferences are included in algorithmic design objectives and outputs are evaluated against these preferences.		
	Disassociated Processing (CT.DP-P): Data processing solutions increase disassociability consistent with the organization's risk strategy to protect individuals' privacy and enable implementation of privacy principles (e.g., data minimization).	**CT.DP-P1:** Data are processed to limit observability and linkability (e.g., data actions take place on local devices, privacy-preserving cryptography).	AC-23	Data Mining Protection
			AU-16(3)	Cross-Organizational Audit Logging | Disassociability
			IA-8(6)	Identification and Authentication (Non-Organizational Users) | Disassociability
			PL-8	Security and Privacy Architecture
			PM-7	Enterprise Architecture
			SA-8(33)	Security and Engineering Privacy Principles | Minimization
			SA-17	Developer Security and Privacy Architecture and Design
		CT.DP-P2: Data are processed to limit the identification of individuals (e.g., de-identification privacy techniques, tokenization).	AC-23	Data Mining Protection
			IA-4(8)	Identifier Management | Pairwise Pseudonymous Identifiers
			SA-8(33)	Security and Engineering Privacy Principles | Minimization
			SI-12(1)	Information Management and Retention | Limit PII Elements
			SI-12(2)	Information Management and Retention | Minimize PII in Testing, Training, and Research
			SI-19	De-Identification
			AC-23	Data Mining Protection

NIST Privacy Framework Core			NIST SP 800-53, Revision 5, Control	
Function	Category	Subcategory	Control ID	Control Name
		CT.DP-P3: Data are processed to limit the formulation of inferences about individuals' behavior or activities (e.g., data processing is decentralized, distributed architectures).	AU-16(3)	Cross-Organizational Audit Logging | Disassociability
			IA-8(6)	Identification and Authentication (Non-Organizational Users) | Disassociability
			PL-8	Security and Privacy Architectures
			PM-7	Enterprise Architecture
			SA-8(33)	Security and Engineering Privacy Principles | Minimization
			SA-17	Developer Security and Privacy Architecture and Design
			SC-2(2)	Separation of System and User Functionality | Disassociability
			SI-19	De-Identification
		CT.DP-P4: System or device configurations permit selective collection or disclosure of data elements.	CM-6	Configuration Settings
			SA-8(33)	Security and Engineering Privacy Principles | Minimization
			SC-42(5)	Sensor Capability and Data | Collection Minimization
		CT.DP-P5: Attribute references are substituted for attribute values.	AC-16	Security and Privacy Attributes
			SA-8(33)	Security and Engineering Privacy Principles | Minimization
	COMMUNICATE-P (CM-P): Develop and implement appropriate activities to enable organizations and individuals to have a reliable understanding and engage in	**Communication Policies, Processes, and Procedures (CM.PO-P):** Policies, processes, and procedures are maintained and used to	**CM.PO-P1:** Transparency policies, processes, and procedures for communicating data processing purposes, practices, and associated privacy risks are established and in place.	PT-1	Policy and Procedures
				PT-2	Authority to Process PII
				PT-3	PII Processing Purposes
				RA-8	Privacy Impact Assessment

NIST Privacy Framework Core			NIST SP 800-53, Revision 5, Control	
Function	Category	Subcategory	Control ID	Control Name
a dialogue about how data are processed and associated privacy risks.	increase transparency of the organization's data processing practices (e.g., purpose, scope, roles and responsibilities in the data processing ecosystem, and management commitment) and associated privacy risks.	**CM.PO-P2:** Roles and responsibilities (e.g., public relations) for communicating data processing purposes, practices, and associated privacy risks are established.	PT-1	Policy and Procedures
	Data Processing Awareness (CM.AW-P): Individuals and organizations have reliable knowledge about data processing practices and associated privacy risks, and effective mechanisms are used and maintained to increase predictability consistent with the organization's risk strategy to protect individuals' privacy.	**CM.AW-P1:** Mechanisms (e.g., notices, internal or public reports) for communicating data processing purposes, practices, associated privacy risks, and options for enabling individuals' data processing preferences and requests are established and in place.	SC-42(4)	Sensor Capability and Data | Notice of Collection
		CM.AW-P2: Mechanisms for obtaining feedback from individuals (e.g., surveys or focus groups) about data processing and associated privacy risks are established and in place.		
		CM.AW-P3: System/product/service design enables data processing visibility.	PL-8	Security and Privacy Architectures
			SA-17	Developer Security and Privacy Architecture and Design

NIST Privacy Framework Core			NIST SP 800-53, Revision 5, Control	
Function	Category	Subcategory	Control ID	Control Name
			SC-42(4)	Sensor Capability and Data | Notice of Collection
		CM.AW-P4: Records of data disclosures and sharing are maintained and can be accessed for review or transmission/disclosure.		
		CM.AW-P5: Data corrections or deletions can be communicated to individuals or organizations (e.g., data sources) in the data processing ecosystem.	SI-18(5)	PII Quality Operations | Notice of Correction or Deletion
		CM.AW-P6: Data provenance and lineage are maintained and can be accessed for review or transmission/disclosure.	AC-16	Security and Privacy Attributes
			SC-16	Transmission of Security and Privacy Attributes
			SR-4	Provenance
		CM.AW-P7: Impacted individuals and organizations are notified about a privacy breach or event.	IR-1	Policy and Procedures
		CM.AW-P8: Individuals are provided with mitigation mechanisms (e.g., credit monitoring, consent withdrawal, data alteration or deletion) to address impacts of problematic data actions.		
PROTECT-P (PR-P): Develop and implement appropriate data processing safeguards.	**Data Protection Policies, Processes, and Procedures (PR.PO-P)**:	**PR.PO-P1:** A baseline configuration of information technology is created and maintained incorporating security	CM-1	Policy and Procedures
			CM-4	Impact Analyses
			CM-6	Configuration Settings

NIST Privacy Framework Core			NIST SP 800-53, Revision 5, Control	
Function	Category	Subcategory	Control ID	Control Name
	Security and privacy policies (e.g., purpose, scope, roles and responsibilities in the data processing ecosystem, and management commitment), processes, and procedures are maintained and used to manage the protection of data.	principles (e.g., concept of least functionality).	CM-7	Least Functionality
			CM-9	Configuration Management Plan
			SA-10	Developer Configuration Management
		PR.PO-P2: Configuration change control processes are established and in place.	CM-4	Impact Analyses
			SA-10	Developer Configuration Management
		PR.PO-P3: Backups of information are conducted, maintained, and tested.	CP-6	Alternate Storage Site
			CP-9	System Backup
		PR.PO-P4: Policy and regulations regarding the physical operating environment for organizational assets are met.	PE-1	Policy and Procedures
		PR.PO-P5: Protection processes are improved.	PL-2	System Security and Privacy Plans
		PR.PO-P6: Effectiveness of protection technologies is shared.	AC-21	Information Sharing
		PR.PO-P7: Response plans (Incident Response and Business Continuity) and recovery plans (Incident Recovery and Disaster Recovery) are established, in place, and managed.		
		PR.PO-P8: Response and recovery plans are tested.		
		PR.PO-P9: Privacy procedures are included in human resources practices		

NIST Privacy Framework Core			NIST SP 800-53, Revision 5, Control	
Function	Category	Subcategory	Control ID	Control Name
		(e.g., deprovisioning, personnel screening).		
		PR.PO-P10: A vulnerability management plan is developed and implemented.	RA-1	Policy and Procedures
			RA-3	Risk Assessment
			RA-5	Vulnerability Monitoring and Scanning
	Identity Management, Authentication, and Access Control (PR.AC-P): Access to data and devices is limited to authorized individuals, processes, and devices, and is managed consistent with the assessed risk of unauthorized access.	**PR.AC-P1:** Identities and credentials are issued, managed, verified, revoked, and audited for authorized individuals, processes, and devices.	IA-1	Policy and Procedures
			IA-2	Identification and Authentication (Organizational Users)
			IA-3	Device Identification and Authentication
			IA-4	Identifier Management
			IA-5	Authenticator Management
			IA-8	Identification and Authentication (Non-Organizational Users)
			IA-9	Service Identification and Authentication
			IA-12	Identity Proofing
		PR.AC-P2: Physical access to data and devices is managed.		
		PR.AC-P3: Remote access is managed.	AC-1	Policy and Procedures
			AC-20	Use of External Systems
			SC-15	Collaborative Computing Devices and Applications
		PR.AC-P4: Access permissions and authorizations are managed, incorporating the principles of least privilege and separation of duties.	AC-1	Policy and Procedures
			AC-2	Account Management
			AC-3	Account Enforcement
			AC-5	Separation of Duties

NIST Privacy Framework Core			NIST SP 800-53, Revision 5, Control	
Function	Category	Subcategory	Control ID	Control Name
			AC-6	Least Privilege
			AC-14	Permitted Actions Without Identification and Authentication
			AC-16	Security and Privacy Attributes
			AC-24	Access Control Decision
		PR.AC-P5: Network integrity is protected (e.g., network segregation, network segmentation).	AC-4	Information Flow Enforcement
			SC-7	Boundary Protection
		PR.AC-P6: Individuals and devices are proofed and bound to credentials, and authenticated commensurate with the risk of the transaction (e.g., individuals' security and privacy risks and other organizational risks).	AC-14	Security and Privacy Attributes
			AC-16	Access Control Decision
			IA-1	Policy and Procedures
			IA-2	Identification and Authentication (Organizational Users)
			IA-3	Device Identification and Authentication
			IA-4	Identifier Management
			IA-5	Authenticator Management
			IA-8	Identification and Authentication (Non-Organizational Users)
			IA-9	Service Identification and Authentication
			IA-12	Identity Proofing
	Data Security (PR.DS-P): Data are managed consistent with the organization's risk	**PR.DS-P1:** Data-at-rest are protected.		
		PR.DS-P2: Data-in-transit are protected.	SC-8	Transmission Confidentiality and Integrity
			SC-11	Trusted Path

NIST Privacy Framework Core			NIST SP 800-53, Revision 5, Control	
Function	Category	Subcategory	Control ID	Control Name
	strategy to protect individuals' privacy and maintain data confidentiality, integrity, and availability.	**PR.DS-P3:** Systems/products/services and associated data are formally managed throughout removal, transfers, and disposition.		
		PR.DS-P4: Adequate capacity to ensure availability is maintained.		
		PR.DS-P5: Protections against data leaks are implemented.	AC-4	Information Flow Enforcement
			AC-5	Separation of Duties
			AC-6	Least Privilege
			PE-19	Information Leakage
			SC-7	Boundary Protection
			SI-4	System Monitoring
		PR.DS-P6: Integrity checking mechanisms are used to verify software, firmware, and information integrity.		
		PR.DS-P7: The development and testing environment(s) are separate from the production environment.		
		PR.DS-P8: Integrity checking mechanisms are used to verify hardware integrity.		
	Maintenance (PR.MA-P): System maintenance and repairs are performed consistent with policies,	**PR.MA-P1:** Maintenance and repair of organizational assets are performed and logged, with approved and controlled tools.		

NIST Privacy Framework Core			NIST SP 800-53, Revision 5, Control	
Function	Category	Subcategory	Control ID	Control Name
	processes, and procedures.	**PR.MA-P2:** Remote maintenance of organizational assets is approved, logged, and performed in a manner that prevents unauthorized access.		
	Protective Technology (PR.PT-P): Technical security solutions are managed to ensure the security and resilience of systems/products/services and associated data, consistent with related policies, processes, procedures, and agreements.	**PR.PT-P1:** Removable media is protected and its use restricted according to policy.		
		PR.PT-P2: The principle of least functionality is incorporated by configuring systems to provide only essential capabilities.	CM-7	Least Functionality
		PR.PT-P3: Communications and control networks are protected.	SC-7	Boundary Protection
			SC-11	Trusted Path
			SC-23	Session Authenticity
			SC-31	Cover Channel Analysis
			SC-38	Operations Security
		PR.PT-P4: Mechanisms (e.g., failsafe, load balancing, hot swap) are implemented to achieve resilience requirements in normal and adverse situations.		

LIST OF REFERENCES

[1] National Institute of Standards and Technology, "NIST Privacy Framework: A Tool for Improving Privacy Through Enterprise Risk Management, Version 1.0," National Institute of Standards and Technology, Gaithersburg, MD, NIST CSWP 01162020, 2020. [Online]. doi: 10.6028/NIST.CSWP.01162020

[2] D. K. Mulligan, C. Koopman, and N. Doty, "Privacy is an essentially contested concept: A multi-dimensional analytic for mapping privacy," *Phil. Trans. R. Soc. A.*, vol. 374, no. 2083, p. 20160118, Dec. 2016. [Online]. doi: 10.1098/rsta.2016.0118

[3] S. Brooks, M. Garcia, N. Lefkovitz, S. Lightman, and E. Nadeau, "An Introduction to Privacy Engineering and Risk Management in Federal Systems," National Institute of Standards and Technology, Gaithersburg, MD, NIST IR 8062, 2017. [Online]. doi: 10.6028/NIST.IR.8062

[4] Caitlin Fairchild, "Pentagon to examine fitness trackers post-Strava," *Nextgov*, Jan. 30, 2018. [Online]. Available: https://www.nextgov.com/analytics-data/2018/01/pentagon-examine-fitness-trackers-post-strava/145591/

[5] Office of Personnel Management, "Cybersecurity incidents," Accessed November 8, 2020. [Online]. Available: https://www.opm.gov/cybersecurity/cybersecurity-incidents/

[6] Joint Task Force Interagency Working Group, "Security and Privacy Controls for Information Systems and Organizations," National Institute of Standards and Technology, Gaithersburg, MD, 2020. [Online]. doi: 10.6028/NIST.SP.800-53r5

[7] SAM.gov, "COVID-19: Proximity tracking program," Jul. 08, 2020. [Online.] Available: https://beta.sam.gov/opp/806be5faa3bb4d51b6b9a25dcb16f9ab/view

[8] "Timeline: Theodore Roosevelt COVID-19 outbreak investigation," *USNI News*, Jun. 23, 2020. [Online]. Available: https://news.usni.org/2020/06/23/timeline-theodore-roosevelt-covid-19-outbreak-investigation.

[9] Department of the Navy, "Department of the Navy Strategy for Data and Analytics Optimization," Washington, DC, USA, 2017. [Online]. Available: https://www.doncio.navy.mil/ContentView.aspx?id=9475

[10] Department of the Navy, "Department of the Navy Information Superiority Vision," Washington, DC, USA, 2020. [Online]. Available: https://www.doncio.navy.mil/ContentView.aspx?id=13181

[11] Kimberly Underwood, “The Navy takes the helm in data management,” *SIGNAL Magazine*, Jan. 27, 2020. [Online]. Available: https://www.afcea.org/content/navy-takes-helm-data-management

[12] Department of the Navy Chief Information Officer, “Jupiter: bringing the power of data analytics to the DON,” *CHIPS*, July-September 2020. [Online]. Available: https://www.doncio.navy.mil/chips/ArticleDetails.aspx?ID=13804

[13] David Carroll, “Artificial Intelligence and Machine Learning for COVID Fleet Readiness,” presented at the AI/ML Coalition of the Willing, Naval Postgraduate School, Nov. 06, 2020.

[14] Executive Office of the President, “Big Data: Seizing Opportunities, Preserving Values,” Washington, DC, USA, 2014. [Online]. Available: https://obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf

[15] President’s Council of Advisors on Science and Technology, “Big Data and Privacy: A Technological Perspective,” Washington, DC, USA, 2014. [Online]. Available: https://bigdatawg.nist.gov/pdf/pcast_big_data_and_privacy_-_may_2014.pdf

[16] 113th Congress, *Federal Information Security Modernization Act of 2014*. 2014, p. 16. [Online]. Available: https://www.congress.gov/113/plaws/publ283/PLAW-113publ283.pdf

[17] *Circular No. A-130*, Office of Management and Budget, Washington, DC, USA, 2016. [Online]. Available: https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/circulars/A130/a130revised.pdf

[18] Executive Office of the President, “Fiscal year 2019–2020 Guidance on Federal Information Security and Privacy Management Requirements,” Washington, DC, USA, 2019. [Online]. Available: https://www.whitehouse.gov/wp-content/uploads/2019/11/M-20-04.pdf

[19] Joint Task Force Transformation Initiative, “Risk Management Framework for Information Systems and Organizations: A System Life Cycle Approach for Security and Privacy,” National Institute of Standards and Technology, Gaithersburg, MD, NIST SP 800-37r2, 2018. [Online]. doi: 10.6028/NIST.SP.800-37r2

[20] National Institute of Standards and Technology, "Framework for Improving Critical Infrastructure Cybersecurity, Version 1.1," National Institute of Standards and Technology, Gaithersburg, MD, NIST CSWP 04162018, 2018. [Online]. doi: 10.6028/NIST.CSWP.04162018

[21] *Risk Management Framework (RMF) for DOD Information Technology (IT)*, DOD Instruction 8510.10, Department of Defense, Washington, DC, USA, 2014. [Online]. Available: https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/851001p.pdf?ver=2019-02-26-101520-300

[22] National Institute of Standards and Technology, "Roadmap for Advancing the NIST Privacy Framework: A Tool for Improving Privacy through Enterprise Risk Management," National Institute of Standards and Technology, Gaithersburg, MD, 2020. [Online]. Available: https://www.nist.gov/system/files/documents/2020/01/16/NIST-Privacy-Framework-Roadmap-v1.0.pdf

[23] *DOD Privacy and Civil Liberties Programs*, DOD Instruction 5400.11, Department of Defense, Washington, DC, USA, 2019. [Online]. Available: https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/540011p.pdf

[24] 107th Congress, *E-Government Act of 2002*. 2002, p. 72. [Online]. Available: https://www.govinfo.gov/content/pkg/PLAW-107publ347/pdf/PLAW-107publ347.pdf

[25] *Department of the Navy Privacy Program*, Secretary of the Navy Instruction 5211.5F, Department of the Navy, Washington, DC, USA, 2019. [Online]. Available: https://www.doncio.navy.mil/ContentView.aspx?id=799

[26] DHA Privacy and Civil Liberties Office, "DHA Privacy Program Plan." Defense Health Agency, Falls Church, VA, USA, 2019. [Online]. Available: https://health.mil/Military-Health-Topics/Privacy-and-Civil-Liberties

[27] Committee on National Security Systems, "Security Categorization and Control Selection for National Security Systems," CNSSI No. 1253, 2014. [Online]. Available: https://www.dcsa.mil/portals/91/documents/ctp/nao/CNSSI_No1253.pdf

[28] L. Ferretti *et al.*, "Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing," *Science*, vol. 368, no. 6491, May 2020. [Online]. doi: 10.1126/science.abb6936.

[29] S. Vaudenay, "Analysis of DP3T: between Scylla and Charybdis," Apr. 2020. [Online]. Available: http://eprint.iacr.org/2020/399

[30] S. Vaudenay, “Centralized or decentralized? The contact tracing dilemma,” May 2020. [Online]. Available: http://eprint.iacr.org/2020/531

[31] Dave Muoio, “Apple, Google’s contact tracing update streamlines user enrollment, asks less of public health developers,” *MobiHealthNews*, September 1, 2020. [Online]. Available: https://www.mobihealthnews.com/news/apple-googles-contact-tracing-update-streamlines-user-enrollment-asks-less-public-health

[32] Bobbie Johnson, “Some prominent exposure apps are slowly rolling back freedoms,” *MIT Technology Review*, Nov. 23, 2020. [Online]. Available: https://www.technologyreview.com/2020/11/23/1012491/contact-tracing-mandatory-singapore-covid-pandemic/

[33] J. Bay *et al.*, “BlueTrace: A privacy-preserving protocol for community-driven contact tracing across borders,” Government Technology Agency, Singapore, 2020. [Online]. Available: https://bluetrace.io/static/bluetrace_whitepaper-938063656596c104632def383eb33b3c.pdf

[34] M. R. Hussein, A. B. Shams, E. H. Apu, K. A. A. Mamun, and M. S. Rahman, “Digital surveillance systems for tracing COVID-19: privacy and security challenges with recommendations,” submitted to ICAICT 2020, Bangladesh, Jun. 30, 2020. [Online]. Available: http://arxiv.org/abs/2007.13182

[35] P. H. Kindt, T. Chakraborty, and S. Chakraborty, “How reliable is smartphone-based electronic contact tracing for COVID-19?,” May 12, 2020. [Online]. Available: http://arxiv.org/abs/2005.05625

[36] Y. Lee, “Singapore rules out Apple, Google’s contact-tracing system,” *Bloomberg Law*, Jun. 16, 2020. [Online]. Available: https://news.bloomberglaw.com/tech-and-telecom-law/singapore-rules-out-using-apple-google-contact-tracing-system

[37] J. Abeler, M. Bäcker, U. Buermeyer, and H. Zillessen, “COVID-19 contact tracing and data protection can go together,” *JMIR Mhealth Uhealth*, vol. 8, no. 4, p. e19359, Apr. 2020. [Online]. doi: 10.2196/19359

[38] N. Eagle and A. Pentland, “Reality mining: Sensing complex social systems,” *Pers Ubiquit Comput*, vol. 10, no. 4, pp. 255–268, May 2006. [Online]. doi: 10.1007/s00779-005-0046-3

[39] DON Privacy Team, “Rules for handling PII by DON contractor support personnel,” *Department of the Navy Chief Information Officer*, Aug. 29, 2018. [Online]. Available: https://www.doncio.navy.mil/ContentView.aspx?id=10719

[40] Y. Sun, J. Zhang, Y. Xiong, and G. Zhu, “Data security and privacy in cloud computing,” *International Journal of Distributed Sensor Networks*, vol. 10, no. 7, p. 190903, Jul. 2014. [Online]. doi: 10.1155/2014/190903

[41] H. Takabi, J. B. D. Joshi, and G.-J. Ahn, “Security and privacy challenges in cloud computing environments,” *IEEE Security Privacy*, vol. 8, no. 6, pp. 24–31, Nov. 2010. [Online]. doi: 10.1109/MSP.2010.186

[42] G. D’Acquisto, J. Domingo-Ferrer, P. Kikiras, V. Torra, Y.-A. de Montjoye, and A. Bourka, “Privacy by design in big data: An overview of privacy enhancing technologies in the era of big data analytics,”European Union Agency for Network and Information Security, Heraklion, Greece, 2015. [Online]. doi: 10.2824/641480

[43] S. L. Garfinkel, “De-identification of Personal Information,” National Institute of Standards and Technology, National Institute of Standards and Technology, Gaithersburg, MD, NIST IR 8053, 2015. doi: 10.6028/NIST.IR.8053

[44] A. Narayanan and V. Shmatikov, “Robust de-anonymization of large sparse datasets,” in *2008 IEEE Symposium on Security and Privacy (sp 2008)*, Oakland, CA, USA, May 2008, pp. 111–125. [Online]. doi: 10.1109/SP.2008.33

[45] Y.-A. de Montjoye, C. A. Hidalgo, M. Verleysen, and V. D. Blondel, “Unique in the crowd: The privacy bounds of human mobility,” *Sci Rep*, vol. 3, Mar. 2013. [Online]. doi: 10.1038/srep01376

[46] N. Shevchenko, T. A. Chick, P. O’Riordan, T. P. Scanlon, and C. Woody, “Threat Modeling: A Summary of Available Methods.” Carnegie Mellon University Software Engineering Institute, Pittsburg, PA, 2018. [Online]. Available: https://resources.sei.cmu.edu/asset_files/WhitePaper/2018_019_001_524597.pdf

[47] LINDDUN, “LINDDUN Privacy Engineering.” Accessed February 14, 2021. [Online]. Available: https://www.linddun.org

[48] A. Gholami and E. Laure, “Advanced cloud privacy threat modeling,” *Computer Science & Information Technology (CS & IT)*, pp. 229–239, Jan. 2016. [Online]. doi: 10.5121/csit.2016.60120.

[49] M. Kayaalp, “Modes of de-identification,” *AMIA Annu Symp Proc*, vol. 2017, pp. 1044–1050, Apr. 2018. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5977668/

[50] S. L. Garfinkel, “De-Identifying Government Datasets,” National Institute of Standards and Technology, Gaithersburg, MD, NIST Special Publication 800-188, 2016. [Online]. Available: https://csrc.nist.gov/publications/detail/sp/800-188/draft#:~:text=De%2Didentification%20removes%20identifying%20information,distributing%20or%20publishing%20government%20data

[51] M. Hintze, “Viewing the GDPR through a de-Identification lens: A tool for clarification and compliance,” *SSRN Journal*, 2016. [Online]. doi: 10.2139/ssrn.2909121

[52] L. Rocher, J. M. Hendrickx, and Y.-A. de Montjoye, “Estimating the success of re-identifications in incomplete datasets using generative models,” *Nature Communications*, vol. 10, no. 1, Art. no. 1, Jul. 2019. [Online]. doi: 10.1038/s41467-019-10933-3

[53] U.S. Department of Health and Human Services , “Methods for de-identification of PHI,” U.S. Department of Health and Human Services, Washington, DC, USA, 2012. [Online]. Available: https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html

[54] J. S. Hiller and R. S. Russell, “Privacy in crises: The NIST privacy framework,” *Journal of Contingencies and Crisis Management*, vol. 25, no. 1, pp. 31–38, 2017. [Online]. doi: 10.1111/1468-5973.12143

[55] Issie Lapowsky, “How cambridge analytica sparked the great privacy awakening,” *Wired*, March 17, 2019. [Online]. Available: https://www.wired.com/story/cambridge-analytica-facebook-privacy-awakening/

[56] J. A. Kroll, N. Kohli, and P. Laskowski “Privacy and policy in polystores: A data management research agenda,” *LNCS: Heterogeneous Data Management, Polystores, and Analytics for Healthcare*, vol. 11721, no. 1, Aug. 2019. [Online]. Available: https://escholarship.org/uc/item/1rq8m88w

[57] International Organization for Standardization, “ISO/IEC 27701:2019,” *ISO.org*. [Online]. Available: http://www.iso.org/standard/71670.html

[58] PEPP-PT, “PEPP-PT.org.” Accessed Septemer 28, 2020. [Online]. Available: https://www.pepp-pt.org

[59] PEPP-PT, “PEPP-PT context and mission.” Accessed September 28, 2020. [Online]. Available: https://404a7c52-a26b-421d-a6c6-96c63f2a159a.filesusr.com/ugd/159fc3_878909ad0691448695346b128c6c9302.pdf

[60] “Building Blocks for Pandemic Management Systems Using Proximity Tracing,” PEPP-PT, 2020. [Online]. Available: https://github.com/pepp-pt/pepp-pt-documentation/blob/master/PEPP-PT-building-blocks.pdf

[61] “Decentralized Privacy-Preserving Proximity Tracing.” The DP-3T Project, May 25, 2020. [Online]. Available: https://github.com/DP-3T/documents/blob/master/DP3T%20White%20Paper.pdf

[62] F. Aisec, “Pandemic contact tracing apps: DP-3T, PEPP-PT NTK, and ROBERT from a privacy perspective,” Cryptology ePrint Archive, Reprot 202/489, 2020. [Online]. Available: http://eprint.iacr.org/2020/489

[63] “Security and Privacy Analysis of the Document ‘PEPP-PT: Data Protection and Information Security Architecture,’” The DP-3T Project, April 19, 2020. [Online]. Available: https://github.com/DP-3T/documents/blob/master/Security%20analysis/PEPP-PT_%20Data%20Protection%20Architechture%20-%20Security%20and%20privacy%20analysis.pdf

[64] “Security and Privacy Analysis of the Document ‘ROBERT: ROBust and privacy-presERving proximity Tracing.’” The DP-3T Project, April 22, 2020. [Online]. Available: https://github.com/DP-3T/documents/blob/master/Security%20analysis/ROBERT%20-%20Security%20and%20privacy%20analysis.pdf

[65] “Data Protection and Information Security Architecture: Illustrated on German Implementation.” PEPP-PT, Apr. 20, 2020. [Online]. Available: https://github.com/pepp-pt/pepp-pt-documentation/blob/master/10-data-protection/PEPP-PT-data-protection-information-security-architecture-Germany.pdf

[66] R. L. Rivest, D. J. Weitzner, L. C. Ivers, I. Soibelman, and M. A. Zissman, “PACT: Private Automated Contact Tracing Mission and Approach,” PACT, Cambridge, MA, 2020. [Online]. Available: https://pact.mit.edu/wp-content/uploads/2020/05/PACT-Mission-and-Approach-2020-05-19-.pdf

[67] Rivest et al., “The PACT Protocol Specification,” PACT, Cambridge, MA, 2020. [Online]. Available: https://pact.mit.edu/wp-content/uploads/2020/04/The-PACT-protocol-specification-ver-0.1.pdf

[68] Apple Inc., “Exposure Notification-FAQ v1.1,” May 2020. [Online]. Available: https://covid19-static.cdn-apple.com/applications/covid19/current/static/contact-tracing/pdf/ExposureNotification-FAQv1.1.pdf

[69] Apple Inc., “Exposure Notification-Cryptography Specification v1.2,” April 2020. [Online]. Available: https://covid19-static.cdn-apple.com/applications/covid19/current/static/contact-tracing/pdf/ExposureNotification-CryptographySpecificationv1.2.pdf?1

[70] H. Cho, D. Ippolito, and Y. W. Yu, “Contact tracing mobile apps for COVID-19: Privacy considerations and related trade-offs,” *arXiv:2003.11511 [cs]*, Mar. 2020. [Online]. Available: http://arxiv.org/abs/2003.11511

[71] National Institute of Standards and Technology, “NIST Privacy Framework and Cybersecurity Framework to NIST Special Publication 800-53, Revision 5 Crosswalk,” National Institute of Standards and Technology, Gaithersburg, MD, 2020. [Online]. Available: https://www.nist.gov/privacy-framework/nist-privacy-framework-and-cybersecurity-framework-nist-special-publication-800-53

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California